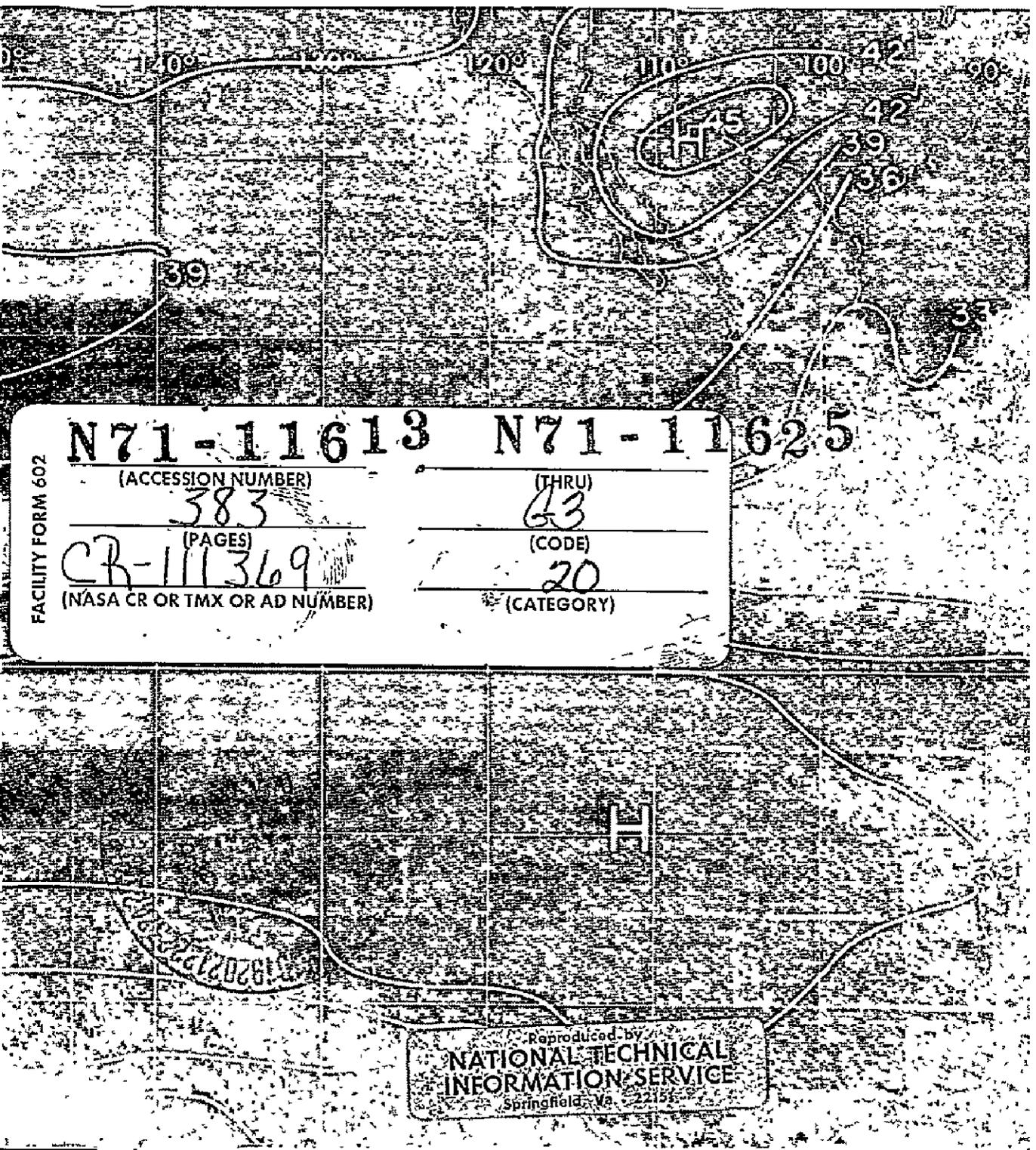


Measurements from Satellite Platforms

Annual Scientific Report on NAS5-11542, 1968-69
Space Science and Engineering Center
The University of Wisconsin
Madison, Wisconsin
August, 1970



FACILITY FORM 602	N71-11613	N71-11625
	(ACCESSION NUMBER)	(THRU)
	383	43
	(PAGES)	(CODE)
CR-111369	20	
(NASA CR OR TMX OR AD NUMBER)	(CATEGORY)	

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

COVER PHOTOGRAPH

This photograph combines the data from two earth orbiting satellites during the period June 16-30, 1969.

The various shades of grey are obtained by photographically averaging daily pictures from the AVCS television camera on the ESSA IX satellite. The resulting picture shows an averaged distribution of clouds over a large portion of the tropics and subtropics.

The line drawing which is superimposed on the photograph is an analysis of the infrared radiation loss from the earth. These measurements were obtained by the medium resolution infrared radiometer on the NIMBUS III satellite; the units are 10^{-2} cal cm^{-2} min^{-1} . Large values of infrared radiation loss are indicative of relatively cloudless areas in which much of the radiation is coming from the warm ocean or land surface. Whereas small infrared radiation loss values are indicative of cloudy areas in which most of the energy is being radiated from the cold cloud tops.

Space Science and Engineering Center
The University of Wisconsin
Madison, Wisconsin

METEOROLOGICAL MEASUREMENTS FROM
SATELLITE PLATFORMS

Annual Scientific Report

on

NAS5-11542

1968-1969

The research in this document has been supported, in whole or in part, by the
National Aeronautics and Space Administration.

August 1970

PRINCIPAL INVESTIGATORS

Verner E. Suomi
Thomas H. Vonder Haar

CONTRIBUTIONS BY

N. H. Beachley
R. S. Cram
R. G. DeDecker
G. J. Dittberner
K. J. Hanson
O. J. Karst
N. Levanon
R. J. Parent

S. K. Peekna
D. H. Sargeant
D. N. Sikdar
A. J. Stamm
A. E. Staver
V. E. Suomi
J. J. Uicker
T. H. Vonder Haar

University of Wisconsin

Technical Monitor: William R. Bandeen

Goddard Space Flight Center, NASA

PREFACE

Research under this contract emphasizes both the meteorological application of measurements from satellites and also engineering studies related to satellites and satellite observing systems. During this first time period under designation NAS5-11542, our work has included study of geosynchronous satellite data as well as that from early and recent polar orbiters. The wide range of topics covered in this document* highlights the imagination and interests of the individual authors. We thank all of them and all others in NASA and The University of Wisconsin for their contributions to our research. We especially thank Miss Betty Ann Erwin for her aid in preparing this report.

Verner E. Suomi
Thomas H. Vonder Haar
Principal Investigators

¹Note that two Special Reports were also issued during 1968-1969. They covered work on the problem of obtaining atmospheric temperature profiles from a geosynchronous satellite and the development of our balloon-borne radio altimeter.

CONTENTS

	Page	
<u>Technical Articles</u>		
1. Geosynchronous Meteorological Satellites, V. E. Suomi and T. H. Vonder Haar	1	✓
2. Applications for Bispectral Radiance Measurements from a Satellite, K. J. Hanson	9	✓
3. Convective Heat Transport over the Tropical Mid-Pacific as Estimated from a Geosynchronous Satellite Altitude, D. N. Sikdar	77	✓
4. A Census of Cloud Systems in the Tropical Pacific Ocean, O. J. Karst	161	✓
5. A Census of Synoptic Scale Disturbances over the Central and Eastern Pacific During March, 1967-February, 1968, A. E. Staver, T. H. Vonder Haar, R. S. Cram, R. G. DeDecker . .	187	✓
6. Possibilities for Quantitative Radiance Measurements in the 450- 650 nm Region from the ATS-I Satellite, S. K. Peekna, R. J. Parent, T. H. Vonder Haar.	201	✓
7. A Study of Cloud Distributions using Reflected Radiance Measurements from the ATS Satellites, A. J. Stamm, T. H. Vonder Haar	219	✓
8. Satellite Measurements of the Earth's Radiation Budget during a 5-year Period, T. H. Vonder Haar.	239	✓
9. A Study of the Indian Monsoon using Satellite Measured Albedo and Long Wave Propagation, G. J. Dittberner	255	✓
10. Simulation and Analysis of Microwave Propagation between Occultation Satellites, D. H. Sargeant	283	✓
11. Balloon-Borne Radio Altimeter, N. Levanon.	347	✓

Contents (continued)

Page

12. Wobble-spin Technique for Spacecraft Inversion and Earth
Photography, N. H. Beachley, J. J. Uicker 379 ✓

N71 - 11614

GEOSYNCHRONOUS METEOROLOGICAL SATELLITE

V. E. Suomi* and T. H. Vonder Haar**
The University of Wisconsin, Madison, Wis.

The great advantage of a geosynchronous meteorological satellite is that the weather moves, not the satellite. It can monitor continuously the weather motions over a large fraction of the earth's surface.

High-resolution photographs from the ATS-I and -III spacecraft (launched in December 1966 and November 1967) have already been used for "weather watches" during hurricane and severe weather conditions. On these spin-stabilized satellites, the spin scan cloud "camera" is really a telescopic photometer whose angle of view (0.1 mrad) is determined by a pinhole in the image plane. Images are found by scanning the scene from west to east using the spacecraft's spin, and scanning from north to south by slowly tilting the telescope through a small angle each revolution. This technique has several advantages. First, there is no optics distortion since all measurements are on-axis and the geometry is generated by the highly uniform spin and tilt. Secondly, the same photometer scans all parts of the picture so the sensitivity is equal everywhere. Thirdly, the detector is a photomultiplier that has a wide dynamic range. Finally, contrast is limited by scatter in the optics only.

In the color camera (Fig. 1) the three color components are generated by using three in-line pinholes. Fiber optics are used to carry the red, blue, and green color information to three separate photomultipliers. The first tube is equipped with a green filter, the second with a red filter, and the third with a blue filter. The signals thus generated are transmitted to the ground. On the next revolution, the telescope is tilted, ever so slightly. After 2400 rev that

Presented as Paper 68-1094 at the AIAA 5th Annual Meeting and Technical Display, Philadelphia, Pa., October 21-24, 1968; submitted October 23, 1968; revision received December 10, 1968. The authors thank the many individuals who contributed to the satellite experiments, especially R. J. Parent.

* Professor of Meteorology and Director, Space Science and Engineering Center.

** Research Scientist, Space Science and Engineering Center.

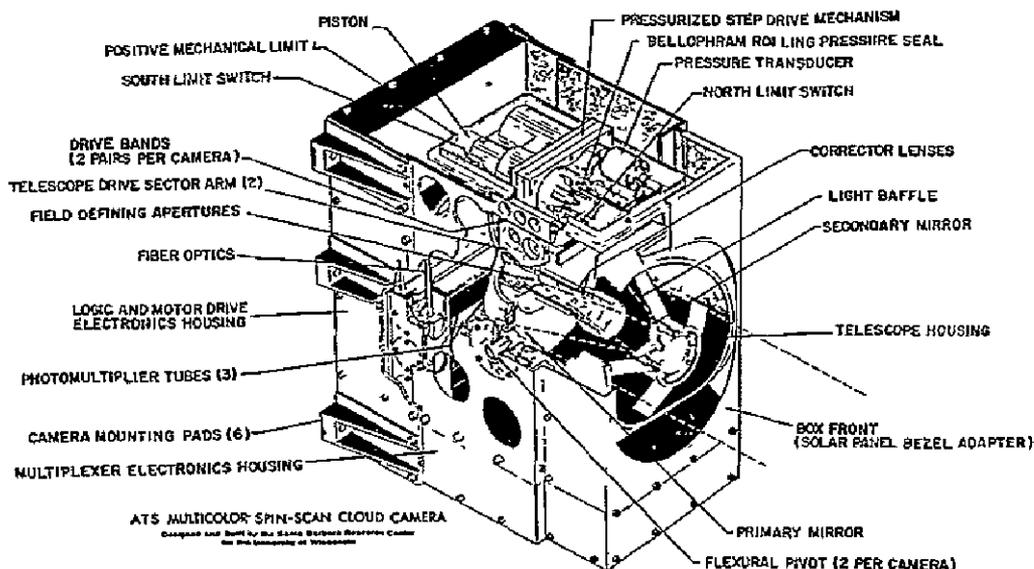


Fig. 1. Multicolor spin-scan cloud camera used to obtain color photographs from NASA's ATS-III geosynchronous satellite.

take about 24 min for spacecraft spinning at 100 rpm, the whole earth is scanned one line at a time from north to south.

With the light signal broken down into three color components, the photomultiplier tubes convert these signals into electrical impulses. The information then is relayed to earth on a single channel using a time-division multiplex technique. The information is sent in groups of four signals. The first signal carries information concerning the intensity of green, the second, red, and the third, blue; the fourth signal is a fixed reference signal to help balance the colors when all the signals are integrated into a complete picture at the receiving stations.

The red and green channels also contain a sun pulse for horizontal synchronization during picture reproduction. Telemetry information provides data for linestepping and vertical synchronization. The vital timing information derived from the sun pulse first locates the camera relative to the sun. Together with ground station input regarding the relative position of earth and sun, the camera scan is then located with regard to the earth. After timing, processing, and other operations, the ATS data are fed to three separate recording systems.

Standard color or black and white negatives are obtained from a photofax recorder able to receive either analog or digital signals and equipped with specially selected, narrow-band color filters. This equipment is located at the Rosman ground station, and photographs are processed at NASA's NIMBUS-ATS Data Center at Goddard Space Flight Center.

Digitized samples of the signal from each channel are also recorded in 8-bit form on magnetic tape. Use of these data allows quantitative studies and displays of the reflected energy from clouds and surface features that utilize the high dynamic range and areal resolution of the cameras. Figure 2, for example, is a brightness analysis of some convective cloud regions in the tropical Pacific. Storing data in this form and using such data in a computer is a major headache, however, since one color photograph in digital form fills three or four reels of high-density magnetic tape. For this reason, application of the digital data usually is confined to detailed studies of portions of a picture or sequence of pictures.

The analog recorder, in contrast, can pack 20 or 30 photos onto one reel, and so ATS data are archived in this form. When we want to use the observations in a computer, an analog-to-digital converter is available to quantize

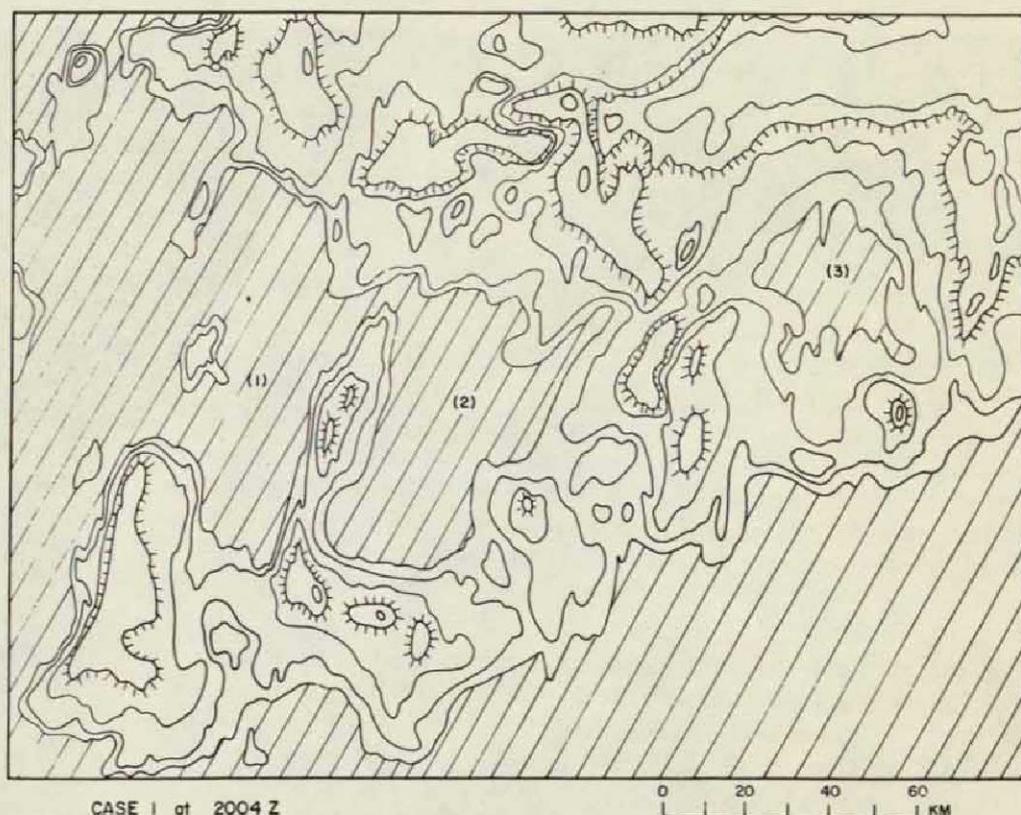


Fig. 2. An analysis of reflected radiance (brightness) values obtained from the ATS-I satellite over a small area of tropical convective clouds. Ocean areas are cross-hatched; brightest cloud areas are set off with "tic" marks.

selected portions of an ATS photo. In addition, playback of the analog tapes through precision electromechanical display devices can produce a photograph of very high quality and resolution (Fig. 3).

Meteorological Application of the ATS Data

When the motion of the weather is speeded up with a time-lapse technique, it is very easy for everyone to see the weather's progress. One can compress into 3 or 4 sec a very good example of diurnal convective activity that occurs over the Amazon basin. As thunderstorm clouds reach high levels, their tops expand horizontally. This daily cycle of cloudiness would make data from sun-synchronous satellites difficult to interpret over this part of the world.

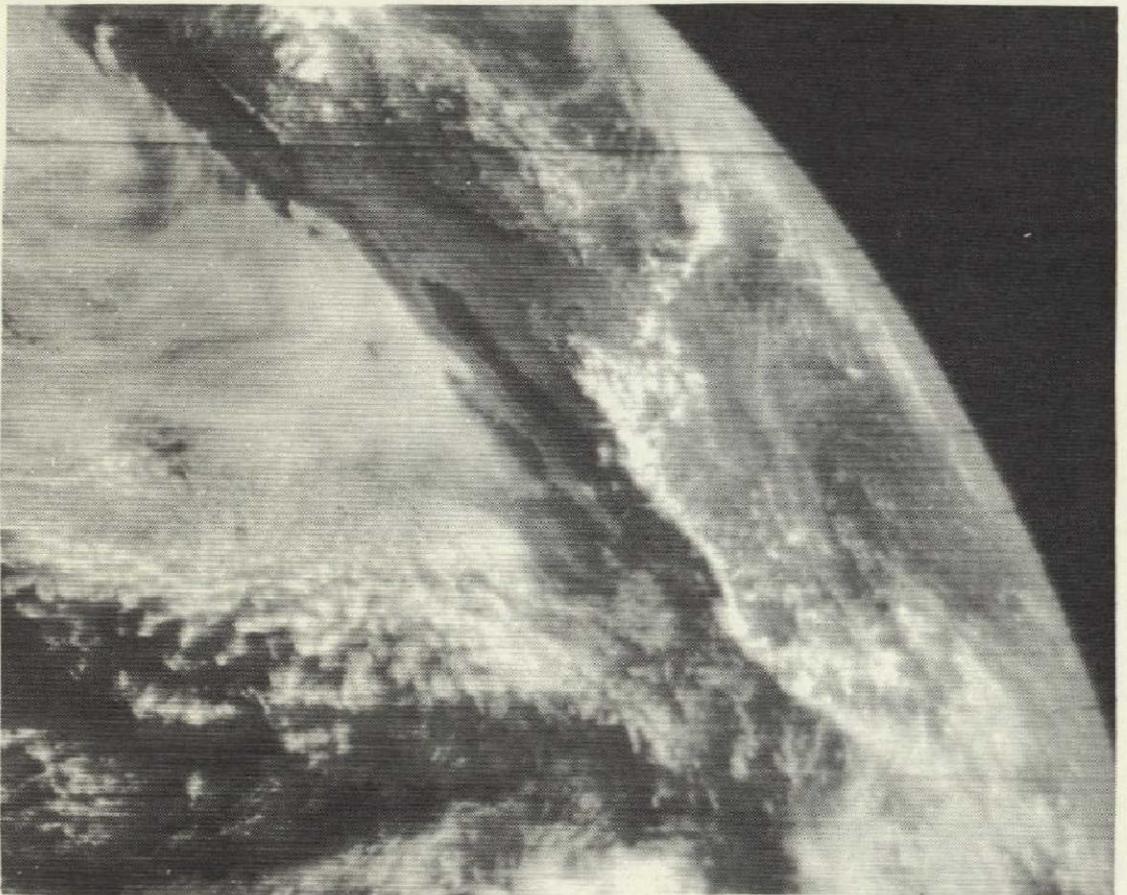


Fig. 3. ATS-I photograph of the Baja California region; coastal stratus clouds and tropical cumulus clouds.

Solar heating over the high Andes and plateaus also produces well-developed convective clouds by late afternoon. None forms over cold Lake Titicaca. A bright white area easily seen in the morning pictures is a large salt flat, 12,000 ft above sea level. This might be a good calibration check for some satellite sensors.

Convective clouds play a vital role in the atmosphere's energy balance by releasing energy into the high troposphere where it can be transported to other regions. In addition, horizontal cloud motions derived from time-lapse photography give information about the wind field that we need to study the direction and variation of this energy transport.

Severe Weather Watch

Violent convective activity over the Midwest is often associated with damaging wind, hail, and tornados. Figure 4 is from a series of ATS pictures obtained every 15 min during a Severe Weather Watch on April 19, 1968. The view covers the eastern two-thirds of the United States, from Florida to Mexico and north slightly beyond the Canadian border. An intense, low-pressure center with its associated cold front was moving towards the Mississippi Valley. Severe weather develops along the warm air (cloudy) side of the front during the day. Near local sunset the damaging "squall line" areas can be clearly seen since they are marked by extensive cirrus blowoff from the thunderstorms. Using the ATS digital data, we can pinpoint the exact location of the thick storm cells near the apex of each plume. At least 10 tornados were associated with these regions of severe weather, and a town in Arkansas was badly damaged by one of them. The real-time integration of pictures such as these into our existing storm warning network will have tremendous value.

Quantitative Measurements from ATS

Several techniques that infer wind velocities and divergence patterns already have been demonstrated to yield reasonable accuracies. Further work to develop fast and economical methods is now underway. Estimates of the vertical energy transport in cumulus clouds can be made by studying the cirrus outflow from them. Preferred organization, size, and life cycles of convective clouds have been measured and will be used to build realistic numerical models. Studies of the scales of organized weather systems in the tropics will help design and locate suitable field programs and provide additional information required for global models of the atmosphere's circulation. Average cloud maps made from the ATS photos have revealed features of the mean atmosphere not found in any textbook.

All of this work has been the result of many research efforts. Most of them will be published soon in Ref. 1, a volume that also will contain papers describing technical details of the ATS Spin Scan Camera experiments.



Fig. 4. An ATS-III photograph of the central U. S. near sundown during a Severe Weather Watch. Brightest clouds within the frontal band mark areas of severe thunderstorm activity. (April 19, 1968)

Summary and Future Prospects

What we have discussed here is only an experiment. It does not yet represent an operation stage. The photos of planet earth that are now available were taken in sunlight during the daytime hours only. The weather motions that we can now observe and study have another drawback; we cannot easily determine the heights of the clouds and thus be sure of the level of the winds we see. Technology already has advanced to a point where suitable infrared radiation detectors make it possible to obtain both daytime and nighttime images. Besides providing continuous images of the cloud motions and locations, it also will be possible to derive the clouds' altitudes from the IR temperature information. In addition, IR measurements in the proper spectral intervals will allow us to sound the atmosphere from a synchronous satellite. While the infrared signal is very low at geosynchronous altitude, we can again use the outstanding attribute of a satellite in such an orbit, its relatively stationary position, to give us a long access time for any region viewed. Resulting temperature profiles are vitally required for worldwide weather

forecasting and for numerical models that simulate atmospheric conditions. Finally, the geosynchronous satellite also can act as the platform from which to radiate a storm warning system. The needed technology is available.

Reference

Suomi, V. and Hanson, K., eds., Weather Motions From Space, University of Wisconsin Press, Madison, Wis., to be published.

N71 - 11615

APPLICATIONS FOR BISPECTRAL RADIANCE MEASUREMENT FROM A SATELLITE

Kirby J. Hanson

ABSTRACT

This study considers the use of Rayleigh scattering as a measure of atmospheric parameters from a geostationary satellite. Two cases are considered: one for clouds with high reflectance; the second for the ocean surface with low reflectance. The objective was to determine (1) cloud top height, and (2) total mass above the sea level surface, or sea level pressure. It is emphasized that the model did not account for particulates.

It was shown that with typical instrument performance parameters, the altitude could be determined to 200 ft and sea level pressure to less than 2 mb, providing a large number of measurements were used. It is essential, in surface pressure determination, that no clouds are in the optical path.

In addition, the study included means to optimize the instrument performance parameters.

Contents

	<u>Page</u>
I. Introduction	13
II. Theory	14
1. Assumed Model Atmosphere	14
2. Atmosphere Scattering Theory	15
a. Radiance Terms	15
b. The Height Discrimination Terms	18
c. Radiance Dependence on $\mu_0, \mu, \phi, A, \tau$	19
1. Dependence on μ	20
2. Dependence on μ_0	24
3. Dependence on A and ϕ	24

III.	Application of the Radiance Inversion Technique	24
1.	Radiance Dependence on Height (Z_0) of the Lower Boundary	25
2.	Sea Level Pressure Determination	31
a.	Pressure Error in P_0	32
3.	Cloud Top Height Determination	35
a.	Height Error in Z_0	35
4.	The Effect of Aerosols and Specular Reflectance	38
5.	Radiance Inversion Technique	39
a.	Results of Iterative Technique	39
b.	Convergence Factors	41
IV.	Tests with Satellite Data	43
1.	Removing Cloud Noise	44
2.	Observed Radiance Distribution and Comparison with Theoretical Radiance Values	53
3.	Test of Inversion Technique on ATS-III Data	57
V.	Possibilities for Improved Pressure and Height Resolution	60
1.	Higher Digitizing Resolution	60
2.	Spectral Separation for Radiance Measurements	60
3.	Large Samples to Reduce Electronic Noise	62
4.	Polarization Measurements	62
VI.	Summary and Recommendations	65
	Appendix—Removing ATS-III Digitizing Noise	66
	References	73

List of Symbols

		<u>First used in</u> <u>Equation #</u>
A	Reflectance of Lower Boundary	1
A/D	Analog to Digital	-
B	Brightness as Observed with the ATS-III SSCC (scale of 0-255)	-

F_0	Sun's Incident Radiance at the Outer Atmosphere	2
F_D	Irradiance at Lower Boundary due only to Direct Solar Beam	2
F_S	Irradiance at Lower Boundary due only to Diffuse Skylight	4
ϕ	Azimuth Angle between Plane of θ_0 and Plane of θ	-
g	Gravitational Acceleration	13
I_C	Radiance through the top of the atmosphere due to Radiance Reflected from the Lower Boundary	9
I_{DD}	Radiance due to Direct Downward and Direct Scattered Upward Radiance	6
I_{DS}	Radiance due to Direct Downward and Scattered Upward Radiance	6
I_e	Stokes' Parameter Parallel Intensity Component	17
$I_{\lambda 1}$	Monochromatic Radiance at Wavelength 1	-
$I_{\lambda 2}$	Monochromatic Radiance at Wavelength 2	-
I_{max}	Maximum Intensity Component	18
I_{min}	Minimum Intensity Component	18
I_S	Radiance Scattered Upward by the Atmosphere	-
I_{SD}	Radiance due to Scattered Downward and Direct Upward Radiance	5
I_{SS}	Radiance due to Scattered Downward and Scattered Upward Radiance	7
I_T	Total Radiance Upward through the Top of the Atmosphere	10
I_{TG}	Radiance (total) Upward through the Top of the Atmosphere in the Green Portion of the Spectrum ($\lambda = 0.535$ microns)	

I_{TB}	Radiance (total) Upward through the Top of the Atmosphere in the Blue Portion of the Spectrum ($\lambda = 0.435$ microns),	-
I_r	Stokes Parameter. Perpendicular Intensity Component	17
K_{λ}^S	Mass Scattering Coefficient	1
L	Local Time at the Subsatellite Point	-
λ	Wavelength of Radiation	1
μ	Cosine of θ	-
μ_0	Cosine of θ_0	-
N	Total Number of Data Points on the Left-hand Side of a Gaussian Curve of Frequency of Occurrence vs. Brightness	-
P	Point of Observation on Earth's Disc	-
P	Polarization of Electromagnetic Radiation	17
P_0	Atmospheric Pressure at the Lower Boundary	13
ρ	Atmospheric Density	1
\bar{s}	A Function of τ	4
T	Chandrasekhar's Transmission Function	6
τ	Rayleigh Optical Thickness	1
θ	Zenith Angle to Observer	3
θ_0	Solar Zenith Angle	2
θ_1	Gaussian Constant	16
θ_2	Gaussian Constant	16
θ_3	Gaussian Constant (Mean Brightness)	16
U	Stokes' Parameter. Determines Orientation of Plane of Polarization	17
z_0	Height of Lower Boundary	1

I. INTRODUCTION

We are now in the period in which preliminary plans are being made for defining instrumentation systems to meet the data requirements for the Global Atmospheric Research Program (GARP U. S. Committee, 1969). In recent years there has been much more attention given to the question of remote sensing by inversion techniques in order to measure atmospheric variables, such as the temperature variation with height (Wark and Hilleary, 1969).

The present study considers the question of how well reflected solar radiation measurements from an earth orbiting satellite can be used in a radiance inversion technique in order to obtain: (1) the heights of cloud tops, and (2) sea level pressure. The basis for this inversion is the theoretical work of Sekera (e. g., 1952 and 1953) in which he obtains the solar radiance reflected from a plane-parallel molecular atmosphere with a diffusely reflecting lower boundary. Of course, the real atmosphere differs from this model in that it contains variable amounts of aerosols, which scatter and absorb as well as absorbing gases, of which O_3 is a principal absorber in the visible portion of the spectrum. Similarly, the actual lower boundary of the atmosphere differs from this diffuse reflectance model in that, if it is a water surface, it exhibits specular (Fresnel) reflectance and, if it is soil or water droplets (clouds), it exhibits bidirectional reflectance properties which are significantly different than diffuse (Lambert) reflectance.

At the present time, theoreticians have not included atmospheric aerosols and specular reflectance at the lower boundary into the equations for reflectance due to multiple scattering from the earth and atmosphere. This is a pressing problem which needs a solution for many applications. One application is that of radiance inversion (to obtain atmospheric parameters) for a realistic model atmosphere.

However, before such an inversion is attempted for an atmosphere with aerosols and realistic reflectance properties, an important question should first be resolved. It is: "If the atmosphere were molecular only, could we measure atmospheric parameters to sufficient accuracy by using a radiance inversion technique, given the constraints of remote sensing by earth satellite, solar illumination angle and photomultiplier noise in the measurement?" This needs to be considered first because, if we cannot make the measurements for a molecular atmosphere, we certainly cannot make them for the real atmosphere. This study, therefore, considers radiance inversion for only a molecular atmosphere with diffuse reflectance at the lower boundary.

II. THEORY

After Lord Rayleigh (1899) discovered the law of molecular scattering which explains the wave length dependence of intensity, the foremost remaining radiation problem in atmospheric physics was that of diffuse reflection and transmission of sunlight in a molecular atmosphere. There was very little progress on this problem until Chandrasekhar (1950) recognized the importance of Stokes representation for its solution. In his book Radiative Transfer, he derived the exact solution for Rayleigh's problem in terms of certain X and Y functions. The solution of these integral equations has been carried out independently at Watson Scientific Computing Laboratory in New York and at the Institute for Numerical Analysis of the National Bureau of Standards located on the campus of the University of California, Los Angeles. From these X and Y functions, the Stokes parameters of the radiation emerging in different directions from the top and bottom of the atmosphere can be computed for different sun elevations and for different reflectivities of the lower boundary. The results of the computations performed at the Western Data Processing Center at the University of California, Los Angeles, are presented by Coulson, et al. (1960). More recently, the equations have been reprogrammed by Dave and Wharten (1968) for general use on high speed computers.

1. Assumed Model Atmosphere

The importance of the work by Chandrasekhar and Sekera to the present study is that it provides tables giving the exact distribution of reflected and transmitted light in a plane-parallel atmosphere, scattering in accordance with Rayleigh's laws.

The assumptions of this model atmosphere (Coulson, 1960) are as follows:

1. Plane-parallel atmosphere in which radiation is scattered according to Rayleigh's law.
2. There is no absorption in the atmosphere; thus the emerging radiation is characterized by the normal optical thickness

$$\tau = \int_{z_0}^{\infty} K_{\lambda}^S(z) \cdot \rho(z) dz \quad (1)$$

where K_{λ}^S is the mass scattering coefficient (which varies by λ^{-4}).

3. The model assumes no atmospheric aerosols.
4. The atmosphere is assumed to be illuminated at the top by parallel unpolarized solar radiation of net flux πF_0 through a unit surface perpendicular to this radiation.

5. Reflectance (A) at the lower boundary is assumed isotropic, unpolarized, and constant for the wavelength increment considered. The model does not allow for transmittance through the lower boundary.

2. Atmospheric Scattering Theory

The radiance emerging outwards from a Rayleigh atmosphere is made up of light scattered from the atmosphere and from the lower boundary surface (clouds, oceans or land). Thus a single measurement of the upward reflected solar radiance at a satellite includes radiation from both of these sources. The present study discriminates between the light scattered upward by the surface and the light scattered upward by the atmospheric column between the satellite and the surface. The latter term is the "signal" studied in the present paper.

a. Radiance Terms

If a plane-parallel molecular atmosphere is illuminated by the sun, the radiance scattered and reflected outward from the atmosphere can be divided conveniently into five components. These components are illustrated in Fig. 1 and will be discussed separately in the following paragraphs. The notation and equations in this section were presented by Coulson (1967) and (1968).

Assume a surface with reflectance A is located at a height Z in a plane-parallel molecular atmosphere and that the sun's incident radiance at the outer atmosphere is F_0 units of energy per unit time, steradian, area and wavelength interval. Also let the solar zenith angle at a point P on the cloud be θ_0 , and let the zenith angle from P to the direction of observation be θ . These angles are shown in Fig. 2. The azimuth angle ϕ is the angular distance from the plane θ_0SC to the plane θOC . Phi (ϕ) is 180 degrees when the sun is behind the observer. We will further define:

$$\mu_0 = \cos \theta_0$$

$$\mu = \cos \theta .$$

The I_{DD} Term. Figure 1 shows a component of radiance which is transmitted directly through the atmosphere to the surface, where it is reflected in the outward direction, θ, ϕ and it passes directly outward through the atmosphere with no scattering. The irradiance incident at the surface is

$$F_D = \pi \cdot F_0 \cdot \cos \theta_0 \cdot e^{-\tau \cdot \sec \phi_0} \quad (2)$$

where τ is the Rayleigh optical thickness of the layer above the cloud. On traversing outward through the atmosphere, the radiance is further attenuated such that

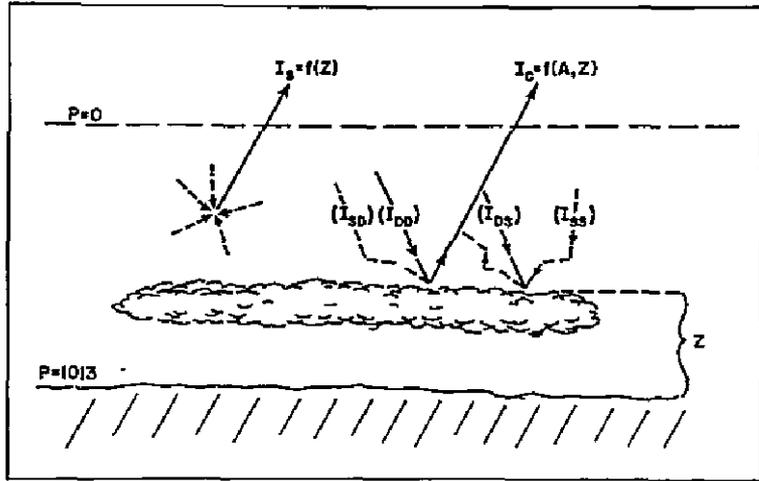


Fig. 1. Radiance Terms for Reflectance from a Rayleigh Atmosphere and a Diffusely Reflecting Lower Boundary.

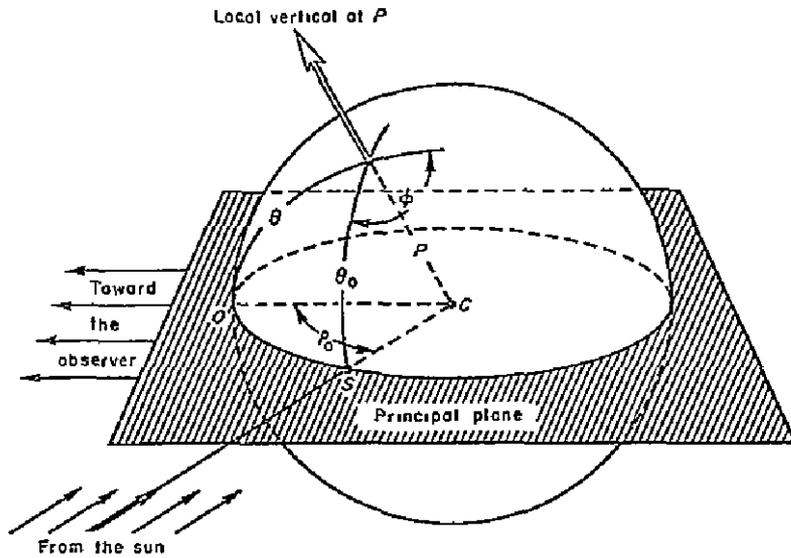


Fig. 2. Coordinates of a Point on a Planetary Disc (from Sekera and Vezee, 1961).

$$I_{DD} = A \cdot F_0 \cdot \cos \theta_0 \cdot e^{-\tau[\sec \theta_0 + \sec \theta]} \quad (3)$$

The I_{SD} Term. The cloud surface is illuminated by diffuse sunlight, part of which is diffuse sunlight transmitted downward from above and part of which is the cloud reflected sunlight which is scattered back down by the overlaying atmosphere. Deirmendjian and Sekera (1954) evaluated the total (i. e. the sum of these two quantities) monochromatic skylight F_S as,

$$F_S = \pi \cdot F_0 \cdot \cos \theta_0 \cdot \left[\frac{\gamma(\theta_0)}{2[1 - A \cdot \bar{s}]} \right] \cdot e^{-\tau \cdot \sec \theta_0} \quad (4)$$

where $\gamma(\theta_0)$ has been computed from Chandrasekhar's X and Y functions and \bar{s} is a known function of only τ . The I_{SD} component is transmitted directly back out through the atmosphere without additional scattering,

$$I_{SD} = A \cdot F_0 \cdot \cos \theta_0 \left[\frac{\gamma(\theta_0)}{2[1 - A \cdot \bar{s}]} \right] \cdot e^{-\tau[\sec \theta_0 + \sec \theta]} \quad (5)$$

The I_{DS} and I_{SS} Terms. After reflection from the cloud top, part of the incident energy is transmitted diffusely upward through the overlaying atmosphere. That is, it is scattered upward through multiple reflections. The two terms considered there concern only that energy which is finally scattered in the direction of observation, θ and ϕ , which denotes an outward direction of measurement. Part of the energy incident on the cloud was transmitted directly downward through the atmosphere, and part of the incident energy is from diffuse skylight. Thus, it is necessary to define two terms which represent this energy. They are given in integral form in terms of Chandrasekhar's (1950) transmission function T as

$$I_{DS} = \frac{A \cdot F_D}{4\pi \cos \theta} \int_0^1 \int_0^{2\pi} T(\tau; \theta, \phi; \theta', \phi') d \cos \theta' d\phi' \quad (6)$$

$$I_{SS} = \frac{A \cdot F_S}{4\pi \cos \theta} \int_0^1 \int_0^{2\pi} T(\tau; \theta, \phi; \theta', \phi') d \cos \theta' d\phi' \quad (7)$$

Here, F_D and F_S are given in equations (2) and (4), respectively. Both of these terms are relatively small and are often computed together in numerical work (Coulson (1967)).

The I_S Term. A part of the radiation emerging from the upper atmosphere will not have reached the cloud top because of scattering in the overlying atmospheric layer. Extensive tables of this term have been computed by Coulson, et al. (1960).

In the present study, the magnitude of the I_g component must be determined in order to evaluate the thickness of the atmosphere overlying the lower boundary—whether this boundary might be a cloud surface or the earth's surface.

b. The Height Discrimination Terms

The previous section described the five radiation terms which define the radiance in an outward direction for a plane-parallel, molecular atmosphere. Numerical values for these terms, as a function of the five physical parameters τ , A , θ_0 , θ , and ϕ have been calculated by previously mentioned investigators from exact theory and are published in various tables; for example, Coulson, et al. (1960).

An indication of the magnitude of the five radiance terms can be seen in Fig. 3 which shows these terms as a function of wavelength (λ) for the case in which

$$\begin{aligned}
 A &= 0.25 \\
 \theta_0 &= 53.1^\circ \\
 \theta &= 50.2^\circ \\
 \phi &= 180^\circ
 \end{aligned}
 \tag{8}$$

The I_g term is largest at short wavelengths and decreases by λ^{-4} . The I_{DD} term, which is direct transmittance downward to the surface and back upward through the atmosphere again, increases with λ because of less attenuation

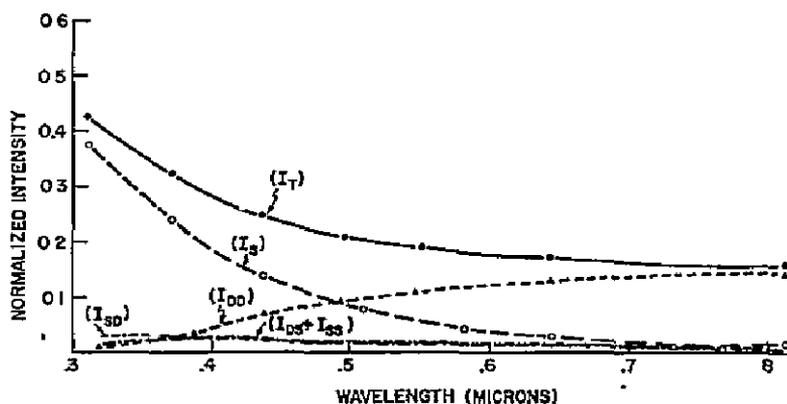


Fig. 3. Radiance Terms vs. wavelength for $A = .25$, $\theta_0 = 53.1^\circ$, $\theta = 50.2^\circ$ and $\phi = 180^\circ$ (from Coulson, 1967).

(by scattering) at longer wavelengths. The three remaining terms, I_{SD} , I_{DS} , and I_{SS} , are relatively small and decrease with wavelength (λ) beyond 0.4 microns because they are primarily scattering terms. The top curve in Fig. 3, I_T , is the sum of the five radiance terms.

Clearly the I_S term is independent of cloud reflectance, whereas the other four terms are all dependent on reflectance as well as atmospheric thickness. Thus, it is convenient to sum the four reflectance dependent terms into a single term I_C ; that is

$$I_C = I_{DD} + I_{SD} + I_{DS} + I_{SS} \quad (9)$$

This makes it possible to express the total radiance I_T as the sum of only two radiance terms

$$I_T = I_S + I_C \quad (10)$$

of which the first is only a function of atmospheric thickness and the second is a function of thickness and reflectance of the lower boundary. That is

$$\begin{aligned} I_S &= f(Z) \\ I_C &= f(A, Z) . \end{aligned} \quad (11)$$

Typical magnitudes for these radiance terms are shown in Fig. 4 ($A = 0.25$) for the same angular condition as Fig. 3.

In summary, there are two height discrimination terms I_S and I_C . The I_S term depends only on the atmospheric mass between the satellite and the lower reflecting surface, whereas the I_C term depends on both atmospheric mass and reflectance of the lower boundary.

c. Radiance Dependence on μ_0 , μ , ϕ , A , and τ

Numerical values of the radiance as a function of the five physical parameters μ_0 , μ , ϕ , A , τ have been calculated previously for the present model atmosphere and are published in various tables; for example, Coulson et al. (1960). This radiance dependence on the five terms is shown in Fig. 5. It is evident that although Coulson's tables are extensive (and occupy a large volume) there are still only a limited number of data points to define the curves. For example, the radiance as a function of reflectance (A) allows only data points at zero, 0.25 and .8. Also, the intensity as a function of μ_0 has only seven data points and the curve is not linear. It is apparent that interpolation over some parts of the curves (such as ϕ , for example) would lead to large errors in radiance. Recently, Dave and Warten (1968) developed a computer program which is a great improvement on Coulson's table in that it provides

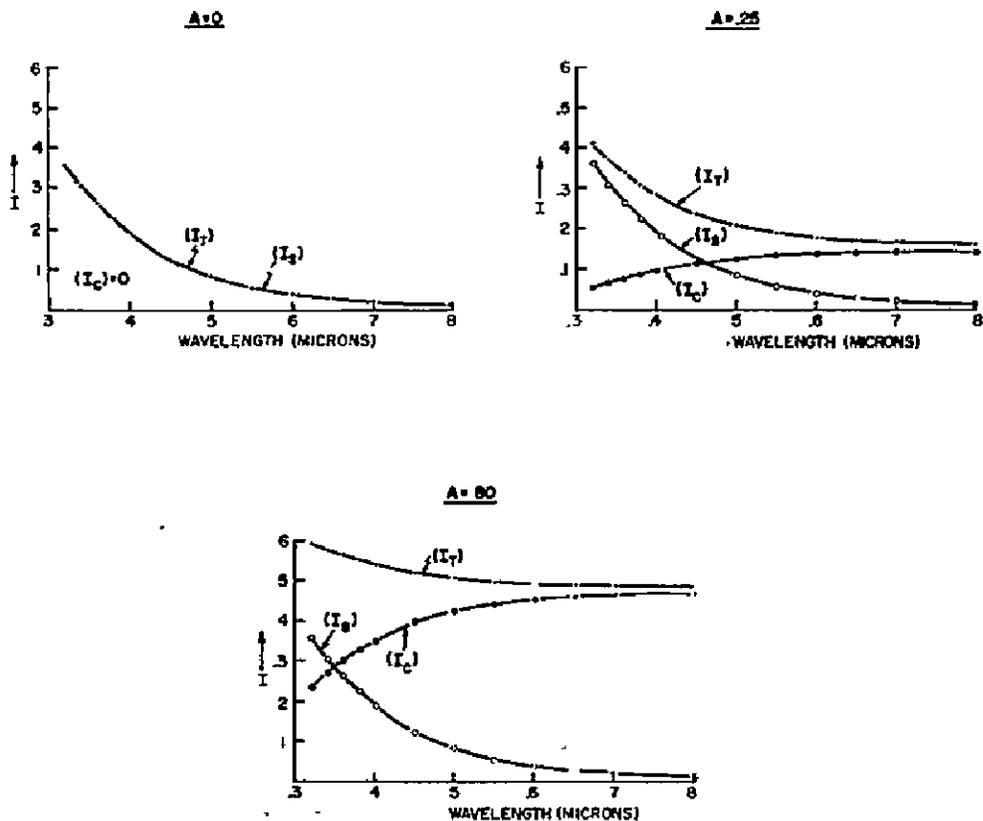


Fig. 4. Radiance Terms for Various Reflectance Values. $\theta_0 = 53.1^\circ$, $\theta = 50.2^\circ$ and $\phi = 180^\circ$.

nearly continuous determination of the radiance (I_T) as a function of the five physical parameters. It actually provides continuous determination of ϕ , τ and A ; and allows for 100 values (between zero and unity) for both μ_0 and μ .

Radiance is the quantity which is fundamental to the present study. Therefore it is important to know and understand its variation due to changes in the five physical parameters. In Figs. 6 and 7 we have illustrated the variation of I_T with four of the five parameters—keeping the last parameter (τ) constant at 0.2238. The basic difference between these two illustrations is that Fig. 6 gives I_T as a function of μ for a range of conditions of A , ϕ and μ_0 . Figure 7, on the other hand, gives I_T as a function of μ_0 for the same range of conditions of A , ϕ and μ .

1. Dependence on μ . The plots in Fig. 6 show that limb brightening ($I_T = f(\mu)$) is common for all angular conditions of μ_0 and ϕ , if the surface

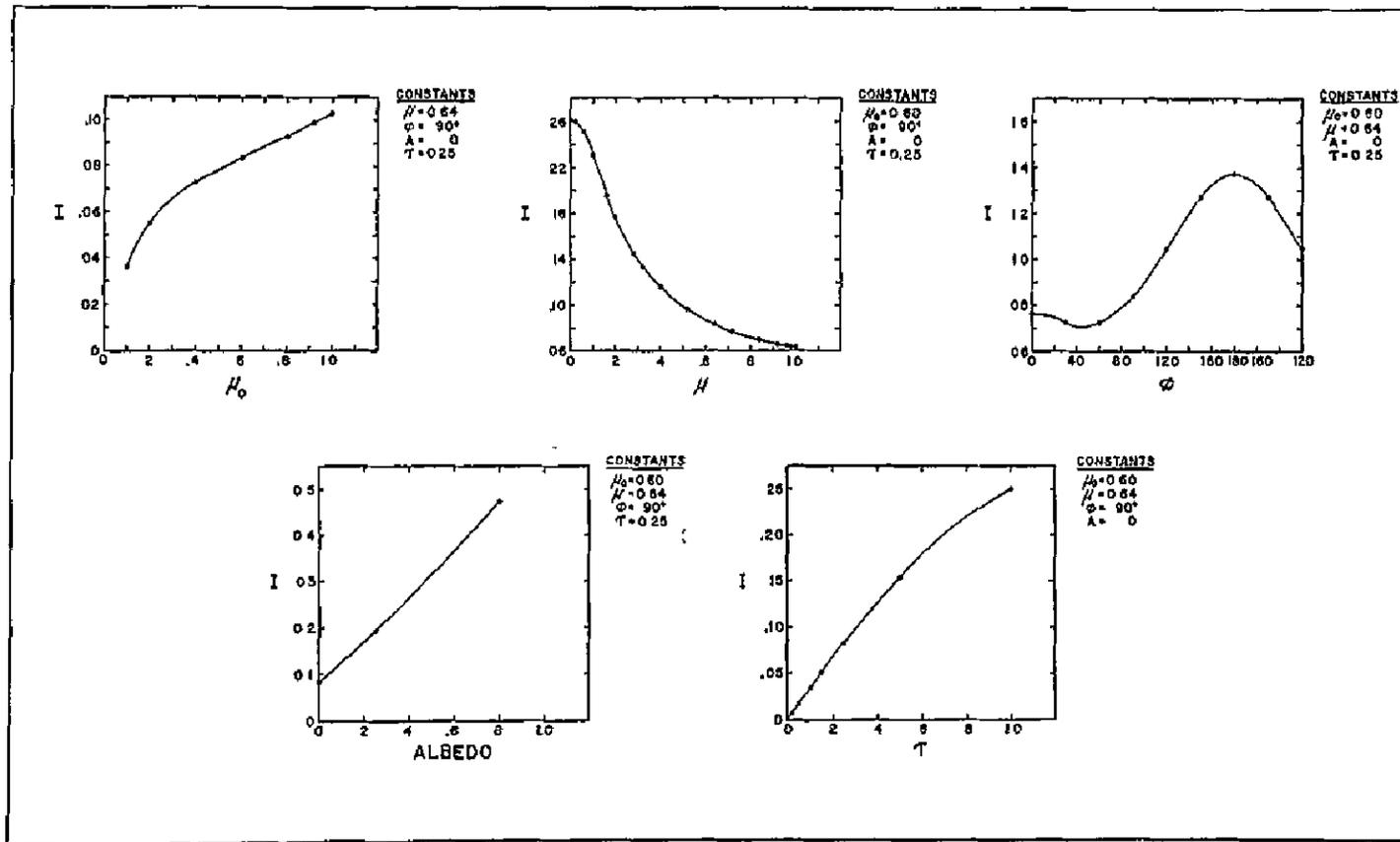


Fig. 5. Radiance Dependence on μ_0 , μ , ϕ , A and τ .

THEORETICAL RADIANCE

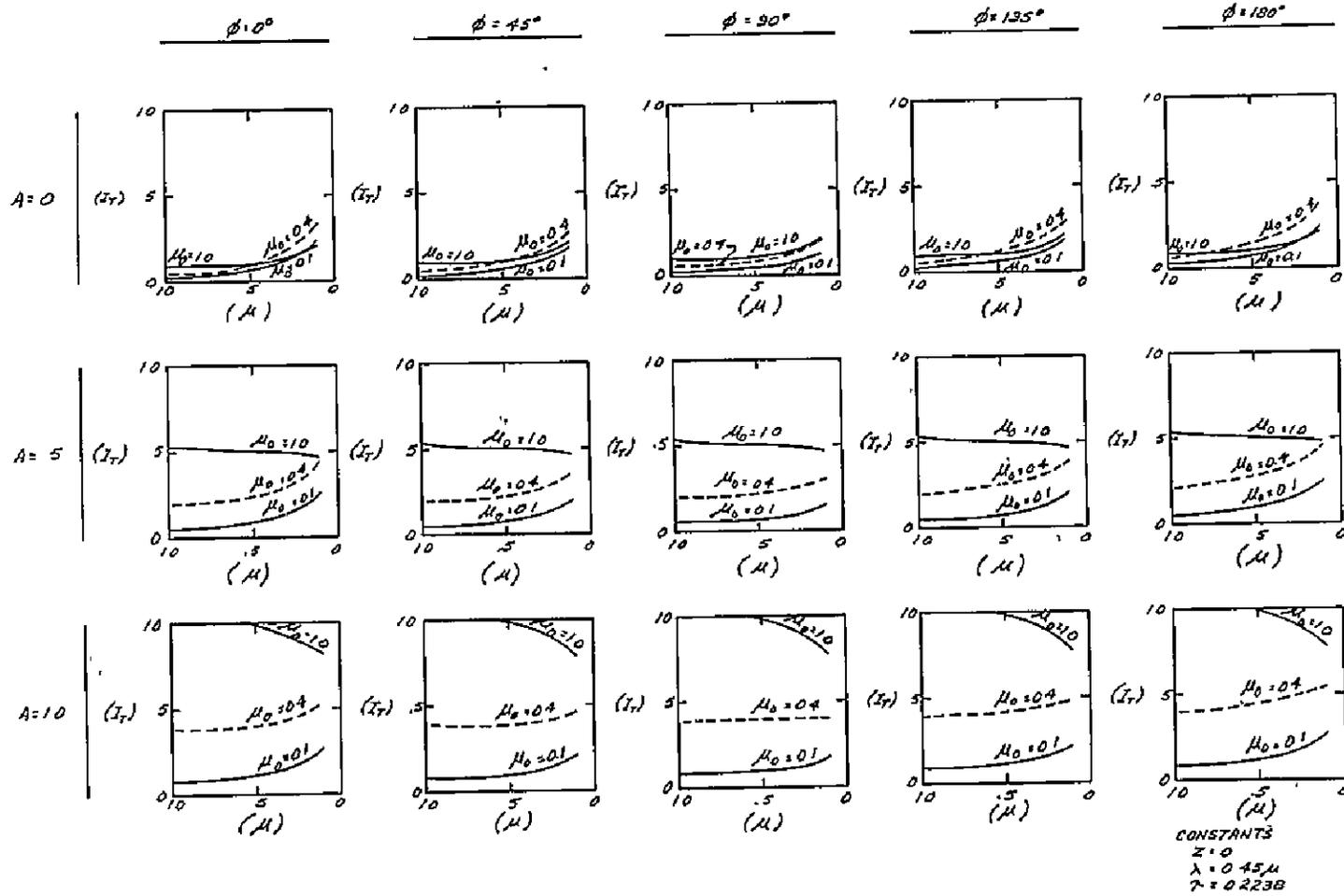
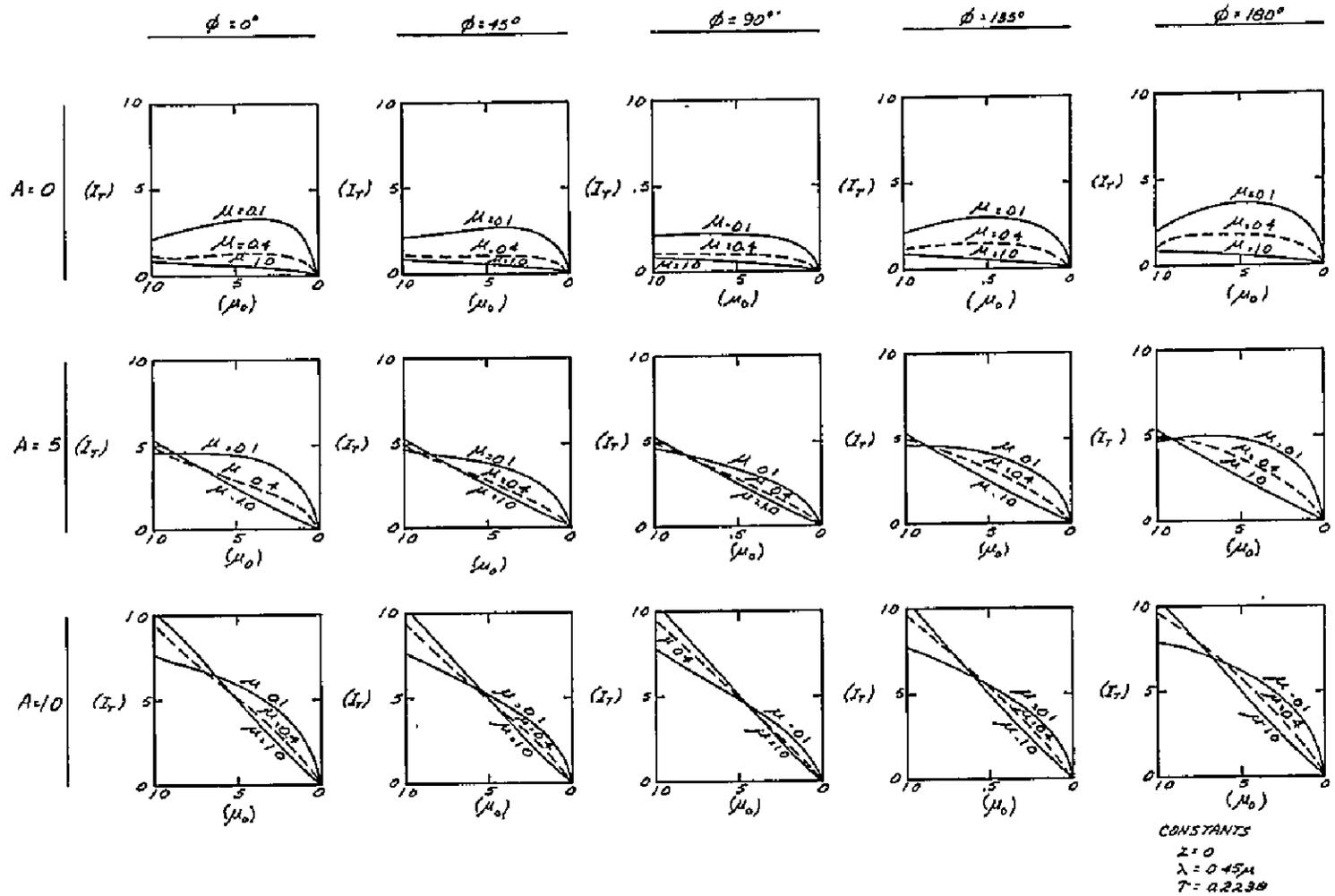


Fig. 6. Radiance Dependence on μ , for certain constant values of μ_0, ϕ, A and τ .

THEORETICAL RADIANCE



23

Fig. 7. Radiance Dependence of μ_0 for certain constant values of μ , ϕ , A and τ .

reflectance is low. However, as the surface reflectance increases, the amount of limb brightening decreases for large μ_0 values and finally becomes limb darkening at large values of A and μ_0 .

Clearly, for the case of a dark ocean and no clouds, limb brightening ($I_T = f(\mu)$) will occur. However, for the case with clouds, there could be either limb brightening or limb darkening, depending on the values of A and μ_0 .

2. Dependence on μ_0 . Figure 7 illustrates how the radiance varies as the sun moves from the zenith to the horizon, for a range of values of A , ϕ and μ .

The top row of data blocks (which is the zero reflectance case) shows that in some cases the radiance increases and in some it decreases, as the sun moves from the zenith to the horizon, depending on the value of μ . We see, for example, that if a satellite sensor is viewing near the subpoint, the radiance decreases systematically as the sun moves from the zenith downward. However, if the satellite is viewing near the limb of the earth the radiance increases until μ_0 is about 0.4 to 0.5 and then decreases toward zero as μ_0 approaches zero. These cases are typical for a low reflectance surface, such as the ocean.

As the reflectance of the surface increases to large values, we see in Fig. 7 that for all values of μ the radiance decreases as the sun moves from the zenith to the horizon. This is the typical case with clouds having a reflectance of $> .50$.

3. Dependence on A and ϕ . We have seen in Figs. 6 and 7 that the radiance depends significantly on A . This has been discussed in the previous two sections. However, we have not considered the dependence of radiance on ϕ . From Figs. 6 and 7 it appears that radiance values are slightly higher when ϕ is either zero or 180° than when $\phi = 90^\circ$. The difference is small in most cases and it is always less than a factor of two.

III. APPLICATIONS OF THE RADIANCE INVERSION TECHNIQUE

In the present study we have considered two applications of the bispectral radiance inversion. One is the determination of cloud top height and the other is the determination of sea level pressure. In principle, both are determined in the same manner, because the radiance inversion yields Z_0 , which is the height of the lower boundary—whether it be the sea surface or a cloud top surface. These two applications are discussed separately in sections 2 and 3 which follow.

The precision by which the inversion determines Z_0 depends on the magnitude of the radiance change with changes in Z_0 . Thus, in order to make estimates of the precision of determining either cloud top height or sea level pressure, it is useful to first consider the magnitude of the radiance changes with changes in Z_0 .

1. Radiance Dependence on Height (Z_0) of the Lower Boundary

In the previous section we have seen how the radiance varied with the four parameters μ_0 , μ , ϕ and A . In that section, the fifth parameter τ was kept constant. This section will show the dependence of radiance on the fifth parameter τ .

Coulson et al. (1960) have defined τ for their calculations as

$$\tau = \int_{Z_0}^{\infty} K_{\lambda}^S(z) \cdot \rho(z) dz \quad (12)$$

This shows clearly that τ depends on wavelength (λ) and the height of the lower boundary (Z_0). The only assumptions needed to determine an explicit relation between τ and Z_0 is to assume density (ρ) and scattering coefficient (K^S) distributions with height (and for the latter value, the variation with wavelength as well). We have used the Rayleigh optical thickness τ vs. Z_0 tables compiled by Elterman (1965) which are based on the U. S. Standard atmosphere. τ vs. Z_0 values from Elterman are given in Table 1 for various wavelengths of radiance.

It is apparent from equation (12) that the dependence of radiance on the height of the lower boundary also depends on the wavelength of the radiance. This is illustrated in Fig. 8 which shows the radiance (I_T) as a function of Z_0 for two wavelengths—0.45 and 0.55 microns. Clearly $\Delta I_T / \Delta Z_0$ is greater at the shortest wavelength (Table 2). It is also evident in Fig. 8 that the change in radiance with change in Z_0 depends on whether a satellite is viewing near the subpoint or near the limb.

To show this angular dependence more clearly, we have illustrated in Figs. 9 and 10 how both I_T and $\Delta I_T / \Delta Z_0$ depend on μ_0 , μ and ϕ . We have selected $\lambda = 0.45$ microns (which is near the peak response of the ATS-III blue channel) to illustrate typical values in the blue portion of the spectrum. However, even larger $\Delta I / \Delta Z_0$ values could be obtained by selecting a shorter wavelength.

Figures 9 and 10 are useful because they show us the sensitivity of the radiance to changes in Z_0 . In Fig. 9, which is for $\phi = 180^\circ$ ($\phi = 0$ is similar), we see that the maximum sensitivity of $\Delta I_T / \Delta Z_0$ occurs with a low

Table 1

Rayleigh Optical Thickness (τ) Values
(from Elterman, 1965)

Height Z_0 (km)	microns			
	0.32	0.45	0.55	1.67
0	0.9290	0.2238	0.0984	0.0011
1	0.8243	0.1986	0.0873	0.0010
2	0.7294	0.1757	0.0772	0.0009
3	0.6435	0.1550	0.0681	0.0008
4	0.5661	0.1364	0.0599	0.0007
5	0.4964	0.1196	0.0526	0.0006
6	0.4338	0.1045	0.0459	0.0005
7	0.3778	0.0910	0.0400	0.0005
8	0.3278	0.0790	0.0347	0.0004
9	0.2833	0.0682	0.0300	0.0003
10	0.2438	0.0587	0.0258	0.0003

reflectance surface ($A = 0$). The maximum value is $-.020[\text{km}^{-1}]$ at $\mu_0 = .6$ and $\mu = .2$; however, for most angular conditions of μ_0 and μ , the sensitivity is between $-.010$ and $-.020$. Note also how the sensitivity diminishes as the surface reflectance is increased. The reason for this appears to be that if a surface has a low reflectance, an upward movement of the lower boundary simply means the (scattering) atmosphere is thinner and there is a decrease in the radiance with increase in Z_0 . However, if the surface has a high reflectance, there is the added factor that a thinner atmosphere allows less attenuation of radiance reflected from the surface. This tends to increase the radiance with increase in Z_0 . Thus, for a high reflectance surface, the increase due to surface radiance offsets the decrease from atmosphere scattering and the net result is that $\Delta I_T / \Delta Z_0$ values are less negative (closer to zero) and are sometimes positive.

Another interesting fact is that although the surface reflectance has a drastic effect on the radiance dependence on μ and μ_0 (see left-hand column of Fig. 9), it has a minimal effect on the change in radiance with Z_0 (see right-hand column of Fig. 9). This change which we see in $\Delta I_T / \Delta Z_0$ with increase in surface reflectance is (1) a decrease in the sensitivity, which has been discussed above, and (2) a gradual shift of the maximum sensitivity toward lower μ_0 values, that is, higher solar zenith angles.

Table 2
 Values of $\Delta I/\Delta Z_0$ (km^{-1})

		A		
		0.05	0.25	0.50
0.36 microns	1.00	.01038	.00629	.00301
	μ 0.64	.02039	.01589	.01176
	0.32	.02540	.02060	.01552

		A		
		0.05	0.25	0.50
0.45 microns	1.00	.0045	.0025	.0012
	μ 0.64	.0099	.0075	.0057
	0.32	.0148	.0119	.0090

		A		
		0.05	0.25	0.50
0.55 microns	1.00	.0021	.0010	.0005
	μ 0.64	.0050	.0036	.0027
	0.32	.0886	.0067	.0051

		A		
		0.05	0.25	0.50
1.67 microns	1.00	.00003	.00001	.00000
	μ 0.64	.00006	.00004	.00006
	0.32	.00011	.00003	.00007

Constants:

$\mu_0 = 0.60$

$\phi = 180^\circ$

$Z = 2 \text{ to } 3 \text{ km}$

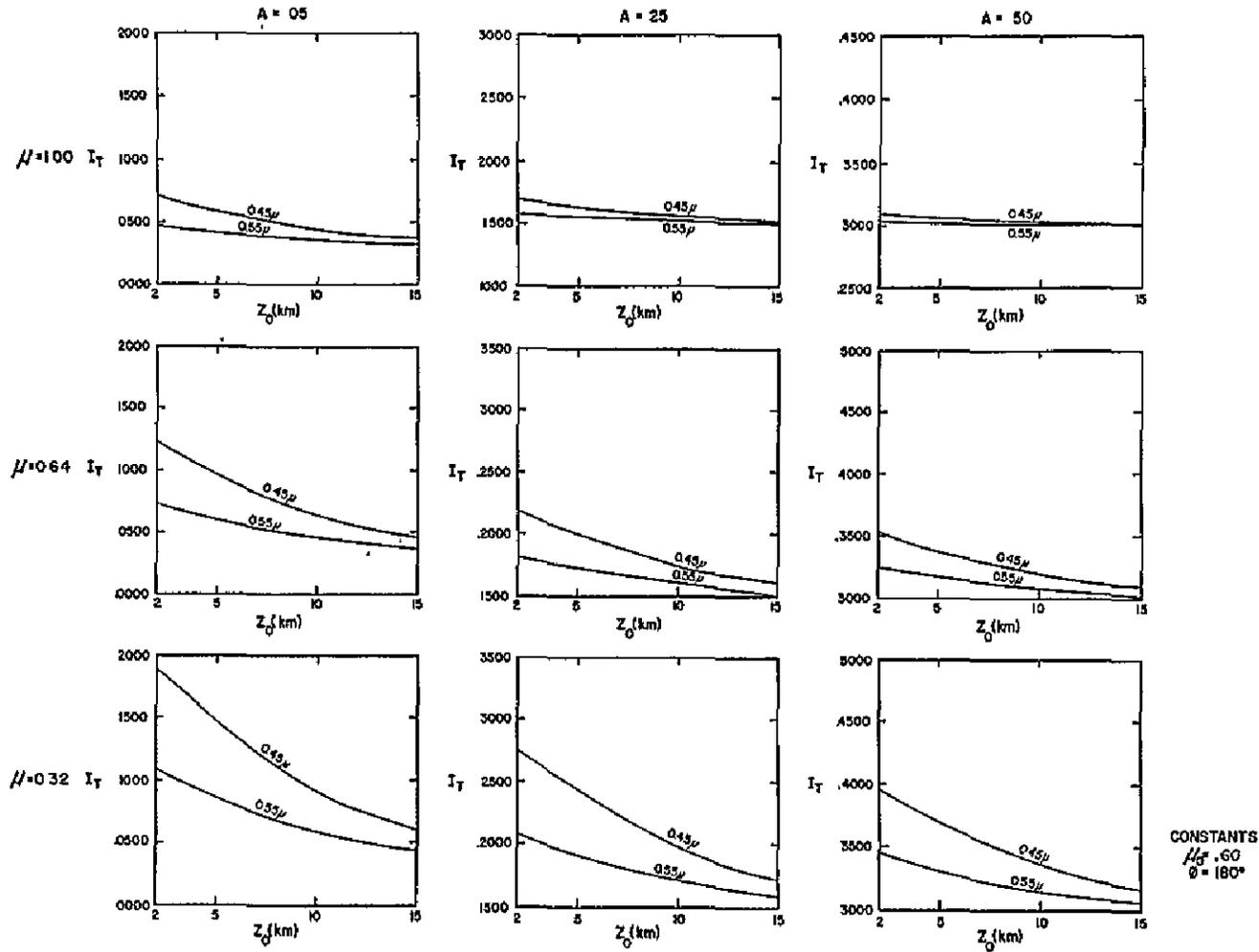


Fig. 8. Radiance Dependence on Z_0 , for certain constant values of μ , ϕ , τ , A , and wavelength values of 0.45 and 0.55 microns.

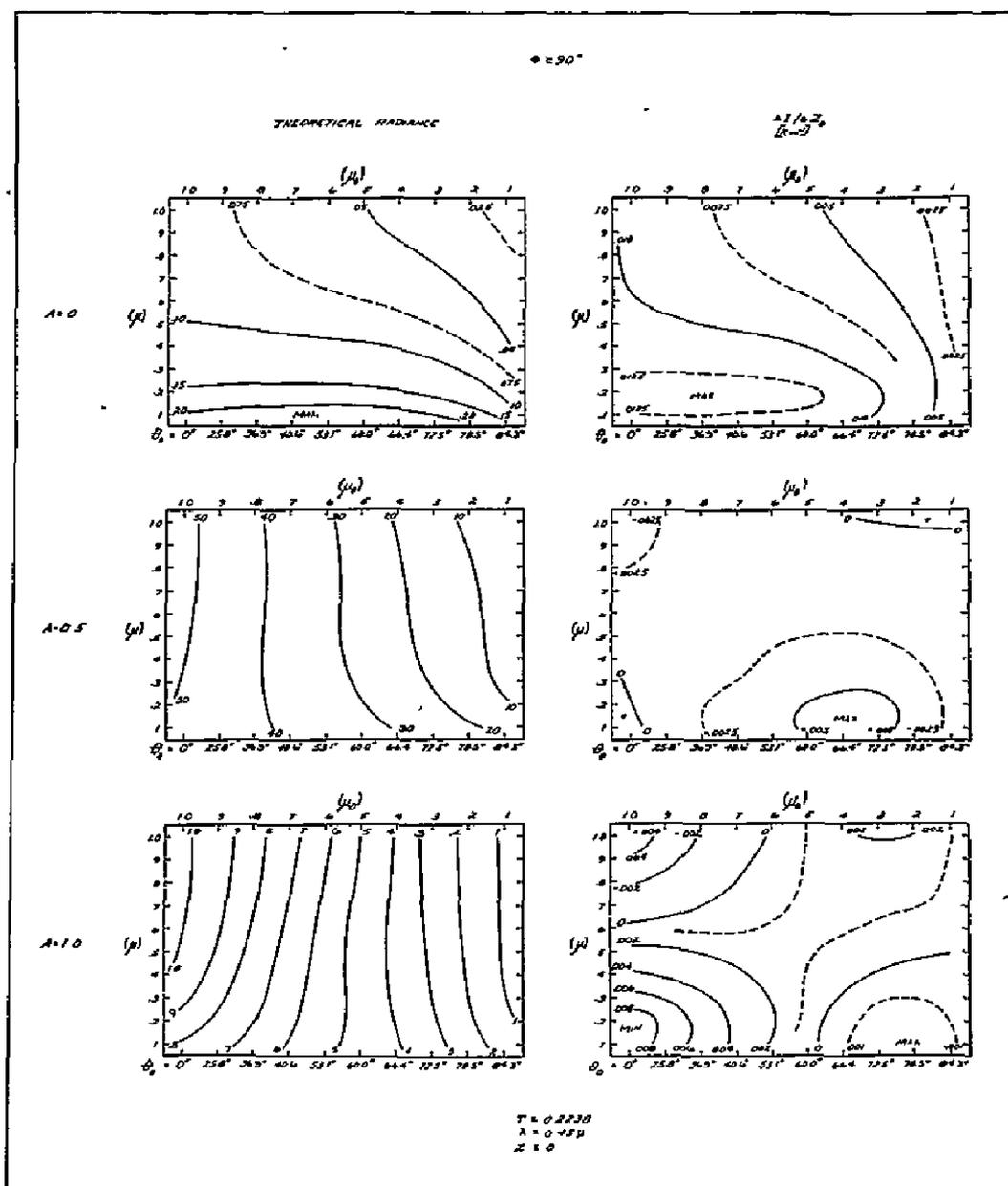


Fig. 10. Dependence of Radiance and $\Delta I/\Delta Z_0$ on μ_0 and μ , for certain constant values of A and for $\tau = .2238$ and $\phi = 90^\circ$.

Another fact of interest is that when the surface has total reflection, maximum sensitivity occurs when μ and μ_0 values are about equal. That is, it is not critical whether their values be large or small, but that they be about the same. This has rather important implications for satellite work—but will be discussed under the section on cloud height errors.

The conditions of Fig. 10 are identical to Fig. 9 except that it represents $\phi = 90^\circ$. Note that although the patterns of radiance and sensitivity dependence on μ and μ_0 are somewhat similar to $\phi = 180^\circ$ (Fig. 9), there is a significant decrease in sensitivity at $\phi = 90^\circ$.

2. Sea Level Pressure Determination

The measurement of sea level pressure on a global scale is needed for GARP because the present network of surface stations covers only a small fraction of the earth (World Weather Program, 1969). The inversion technique being studied here applies in principle to the determination of sea level pressure. This is apparent if one combines the hydrostatic equation and equation (12) and obtains,

$$\tau = -\frac{1}{g} \int_{P_0}^0 K_{\lambda}^S(p) dp \quad (13)$$

This shows that an explicit relationship between τ and P_0 is obtained if we assume a distribution of K_{λ}^S with pressure (p). The gravitational constant (g) is not strictly a constant over the range from P_0 to the P value at satellite altitude. However, it is known quite precisely and could be evaluated within the integral.

A logical first question which should be answered concerning pressure measurement is, for the model atmosphere and surface condition assumed, whether we have reason to believe that pressure measurements can be obtained to sufficient accuracy for the various conditions imposed by satellite measurement (the angular and photomultiplier signal to noise (S/N) conditions). Thus, the question which the present study addresses itself to is, "What pressure errors would be expected with a plane-parallel, homogeneous Rayleigh atmosphere, with a diffuse reflector at the lower boundary, and with measurement obtained by a photomultiplier (PM) tube on an earth synchronous spacecraft?" This question is answered in the next section.

It is important to understand, however, that there are additional questions which are not considered in the present work. These are, "What pressure errors would result from Fresnel reflectance at the sea surface, what pressure errors would result from a non-Rayleigh atmosphere and what pressure errors would result from the presence of clouds which cannot be resolved by the areal resolution of the satellite sensor?"

a. Pressure Error in P_0 . Typical conditions for determining sea level pressure were assumed in order to estimate pressure errors. We have assumed a sea level pressure of 1013 mb, a wavelength of 0.45 microns. The resulting τ value is 0.2238. We have also assumed a surface reflectance of 0.10, based on the data of Krinov (1947).* For these conditions, we have calculated (shown in Fig. 11) how the radiance (I) and the radiance sensitivity ($\Delta I/\Delta Z_0$) vary for all values of μ_0 and μ . As would be expected, they are nearly the same as in $A = 0$ in Fig. 9--differing only because the reflectance is slightly higher in this assumed case.

Also shown in Fig. 11 is the pressure error that results from electrical noise of a photomultiplier tube (PM) (such as the SSCC on ATS-III)** and from $\Delta I/\Delta Z_0$, for all angular conditions of μ_0 and μ . This shows that the gradient in pressure error is reasonably small over nearly all angular conditions. The error is < 200 mb for most angular conditions. This error (for a single measurement) is large; however, for ATS-type data an area of 460 km on a side near the subsatellite point contains 30,000 radiance measurements. Near the earth's limb this number of measurements covers a much larger surface area, of course. If we assume one-third of these measurements contain useful information on surface brightness, then the 10,000 measurements will reduce the pressure error due to random electrical noise by a factor of 100, because random noise is reduced by $1/\sqrt{n}$, where n is the number of measurements. Thus, the minimum pressure error (in Fig. 11) would be .78 mb, and for most angular conditions, the pressure error would be less than 2 mb, for this model atmosphere.

Pressure errors for all angular values of μ_0 and μ are shown in Fig. 11. For a synchronous satellite, however, a particular combination of μ_0 and μ values will occur for points on the earth's disc, depending only on the time of measurement. Thus, for a given point on the earth, the pressure error due to electrical PM noise will depend solely on the time of measurement. We have illustrated these pressure errors in Fig. 12, in which it is seen that the optimum measurement time is at local noon of the subsatellite point. The error is largest at that point at noon and smallest toward (but not at) the earth's limb. The pressure error gradient is quite flat and only varies by a factor of about 2. At 0600 and 0900 local time at the satellite subpoint, it is apparent that pressure errors are considerably larger--especially toward the terminator,*** Along the earth's limb opposite the terminator, the pressure error remains about the same as for that same location at local noon.

*Cited by Elterman and Toolin, 1965 (Fig. 7-4).

**ATS-III Spin-Scan Cloud Camera (SSCC) PM noise characteristics were provided by R. Parent, 1969. The noise is proportional to the square root of the PM output signal.

***Errors due to the plane-parallel approximation become increasingly larger approaching the terminator.

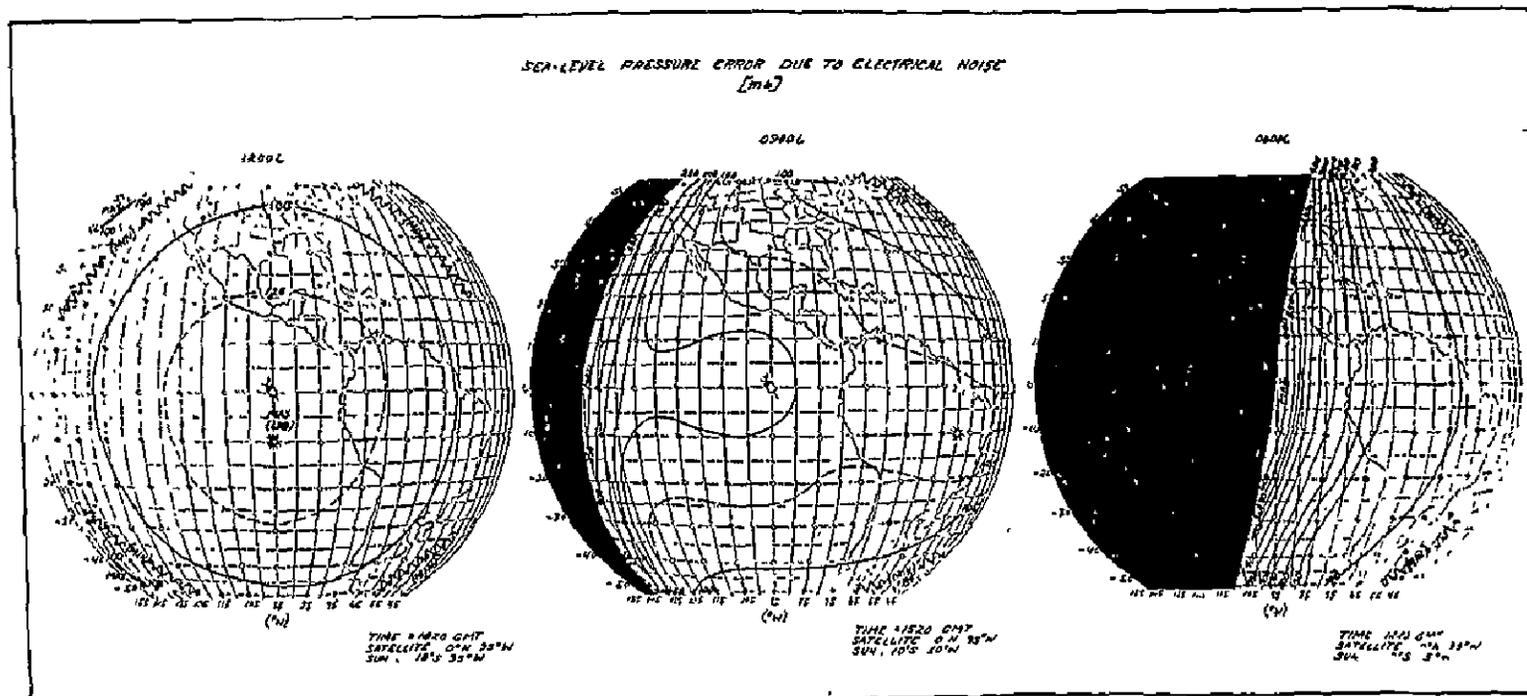


Fig. 12. Sea Level Pressure Error Due to Electrical Noise for Observation Times at 0600, 0900 and 1200 Local Time at the Subsattellite Point.

Again, these pressure errors (in Fig. 12) may be reduced by a factor of 100 if mean sea level pressures over 460×460 km areas are determined. Thus, for all the earth's disc at 1200 local time, the sea level pressure error due to electrical noise of the PM tube is between 0.8 and 1.4 mb. This error is for the assumed model atmosphere with aerosols and Fresnel reflectance at the lower boundary.

3. Cloud Top Height Determination

Cloud top heights can be determined by measurement of IR radiance in "window" portions of spectrum and by a knowledge of the vertical temperature distribution in the atmosphere. Others, such as Wark (1965), have attempted more direct methods of determining cloud top height by measuring the O_2 absorption in an atmospheric column between the cloud and the observer as a measure of its thickness. Errors of several hundred millibars were found in determining cloud top heights. These errors result from the fact that clouds do not have the same (simple) properties which must be assumed for the models. These errors are discussed by Saiedy *et al.* (1965).

We have seen previously from equation (12) that an inversion of radiance yields Z_0 . This section answers the question of "what error is expected in cloud-top height due to photomultiplier electrical noise with the model atmosphere assumed and with the conditions observed by synchronous satellite?" Because of the model atmosphere assumed, the "cloud" top is a horizontally homogeneous bright surface which has diffuse reflectance properties and a transmittance of zero. Thus, it differs from clouds in the real atmosphere which are not characterized as Lambertian surfaces at a well defined height, Z_0 (Bartman, 1967).

a. Height Error in Z_0

In order to determine the height error in Z_0 , we have assumed typical conditions for a "cloud" case. We have let the reflectance at the "cloud" surface (Z_0) be 0.50, the height of Z_0 be 2 km, and the wavelength be 0.45 micron. This gives a Rayleigh optical thickness (τ) of 0.1757. For these conditions we have calculated (and shown in Fig. 13) the radiance (I) and the radiance sensitivity ($\Delta I/\Delta Z_0$).

The radiance sensitivity in the "cloud" height case is less than that for the sea level pressure case, because the reflectance is higher. In fact, sensitivity values near zero are seen in Fig. 13 with the combination of small μ and large μ_0 values—and vice versa. Obviously it is not possible to determine "cloud" heights under these angular conditions in which $\Delta I/\Delta Z_0$ is close to zero.

Height errors resulting from PM electrical noise are shown in the lower part of Fig. 13. We see an increase in height error for the angular conditions mentioned above; however, for other conditions the error gradient is reasonably flat.

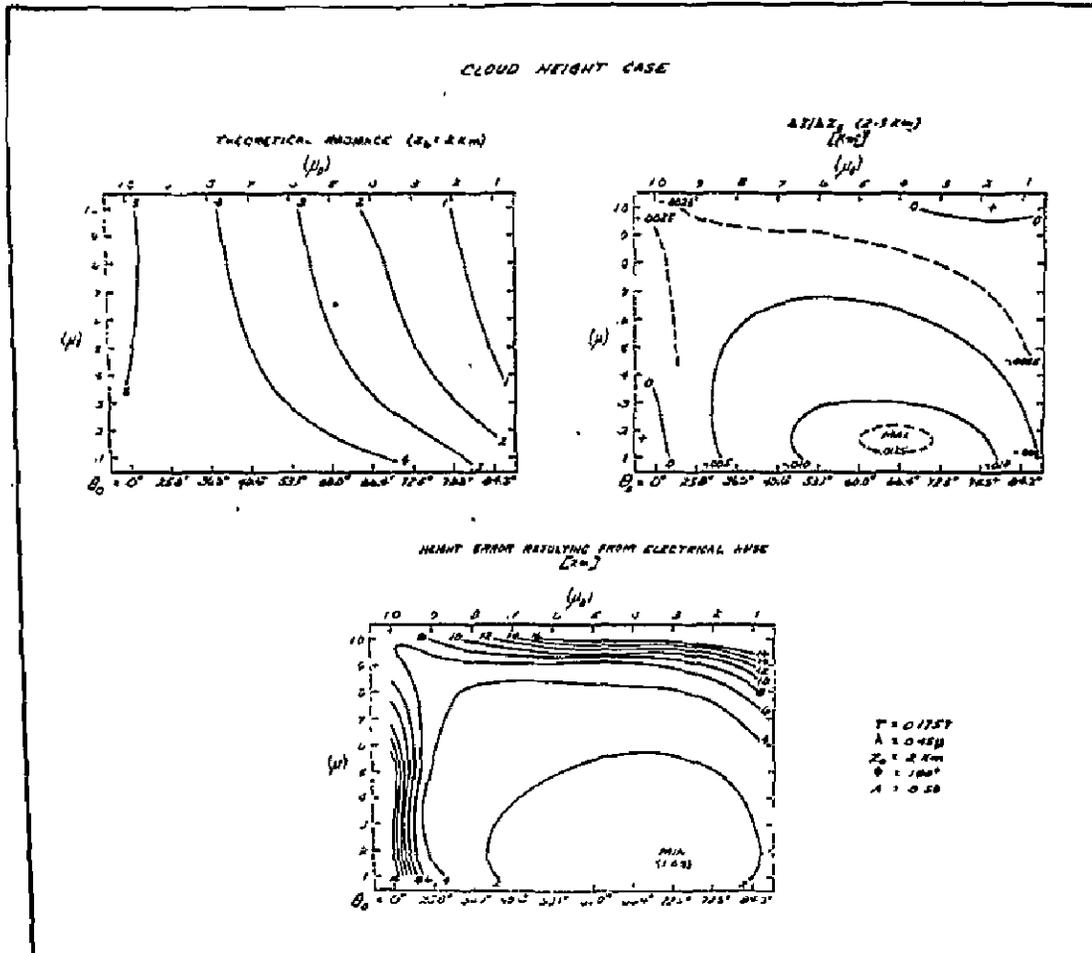


Fig. 13. Dependence of Radiance, $\Delta I / \Delta Z_0$ and Height Error on μ_0 and μ , for $\tau = .1757$, $\phi = 180^\circ$ and $A = .50$.

CLOUD HEIGHT ERROR DUE TO ELECTRICAL NOISE
[km]

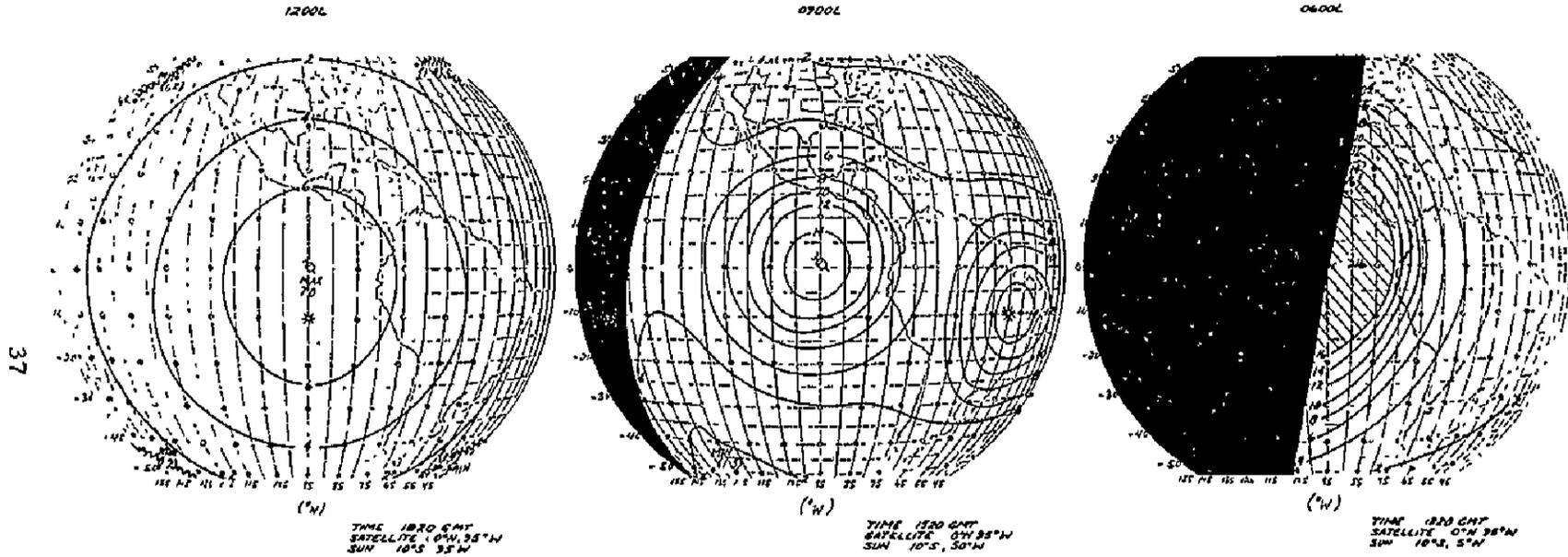


Fig. 14. Cloud-Height Error Due to Electrical Noise for Observation Times at 0600, 0900 and 1200 Local Time at the Subsattelite Point.

For synchronous satellite applications, the height error will depend on the location of the "cloud" on the earth's disc and the time of measurement. "Cloud" height errors of 0600, 0900 and 1200 local time at the satellite subpoint are illustrated in Fig. 14. The units are kilometers (km). As we saw previously for sea level pressure, the optimum time for measurement is at 1200 local. At this time the error is maximum at the subpoint and is greater by a factor of 6 than errors near the earth's limb. (For pressure it was greater by a factor of only 2.) At 0600 and 0900, the error in Z_0 is considerably worse at the subsolar and subsatellite points.

Although this "cloud" height error is large based on a single measurement, it can be reduced considerably for large "clouds" represented by many measurements. For example, if a "cloud" is 30 km on a side, the height error due to random PM noise can be reduced by a factor of 17 if scanned by a sensor having a 3 km ground resolution. Thus, the height errors at 1200L (Fig. 14) for these size "clouds" would range from about 200 to 1100 ft. over the earth's disc.

4. The Effect of Aerosols and Specular Reflectance

The error analysis of the previous two sections have assumed a molecular model atmosphere and a diffusely reflecting surface. Further studies will ultimately have to show the error analysis for an atmosphere with aerosols and gaseous absorbers and for realistic surface reflectance models. At present it is possible to evaluate the effects of these changes in the model only qualitatively—because the theoretical solution has not been completed for multiple scattering with these improved models.

Studies by Shettle and Weinman (1969) show that if one assumes a relatively simple model (primary scattering, and a simple scattering function for aerosols) it is possible to estimate the relative effect of Rayleigh vs. aerosol scattered irradiance outside the atmosphere. Although the effect of aerosols is relatively small when μ_0 is large and the wavelength (λ) is near 0.40, the relative effect of aerosols increases with small values of μ_0 and large values of λ . At $\lambda = .7$ micron, and $\mu_0 = .3$, for example, the radiance from aerosols appears to be about 3 times greater than that from molecular scattering. This fact may eventually serve as the basis for a determination of aerosol concentration—which could then be used to correct for the radiance scattered from these relatively large particles.

Specular reflection from the sea surface also departs from the diffuse reflectance model which we have assumed. The model underestimates the radiance within the sun glint area and overestimates the radiance outside the sun glint area. This is shown clearly in Fig. 22 and 28 which illustrate observed and model calculated radiance values. It is difficult to determine qualitatively

how specular reflection affects the radiance terms because it changes both reflectance from the surface and the effective path length upward through the atmosphere.

The accuracy of radiance inversion also depends on the effect of radiance from clouds which only partially cover the field of view of a sensor or in some other manner act as "noise" to the intended measurement. The elimination of cloud "noise" is discussed in Chapter IV.

5. Radiance Inversion Technique

We have seen previously that it is possible to determine radiance as a function of the five physical parameters assuming a plane-parallel molecular atmosphere.

$$I_{\lambda} = f(\mu_0, \mu, \phi, A, \tau) \quad (14)$$

Now assume in equation (14) that μ_0 , μ and ϕ are known. Under that condition, radiance (I) is only a function of A and τ , which implies that an inversion may be obtained on equation (14) such that two measurements of the radiance at different wavelengths ($I_{\lambda 1}$, $I_{\lambda 2}$) will yield values of A and τ .

The analytical solution to the radiance inversion is being completed by Sekera (1969) at UCLA at the present time. He has previously completed the inversion of the polarization characteristics of radiation emerging from a planetary atmosphere in order to evaluate optical thickness (Sekera, 1967).

In the absence of the analytical solution of the radiance inversion, it is possible to obtain an iterative solution of the inversion. An inversion of this type has been developed in the present study and is shown in Fig. 15. In this inversion the radiance scattered from the atmosphere (I_S) is evaluated at the shortest wavelength and the radiance scattered from the surface and atmosphere (I_C) is evaluated at the longest wavelength. Therefore, τ is estimated from I_S and A is estimated from I_C .

a. Results of Iterative Technique

In employing this technique we found it is possible to begin with values of radiance of two wavelengths, assume $A = 0$ and $Z = 0$, and, through the above iteration procedure, to converge on proper values of A and τ .

We find this iteration scheme converges logarithmically to the correct height* and reflectance. A typical example is shown in Fig. 16. We also find

*Height may be used instead of τ because λ is shown.

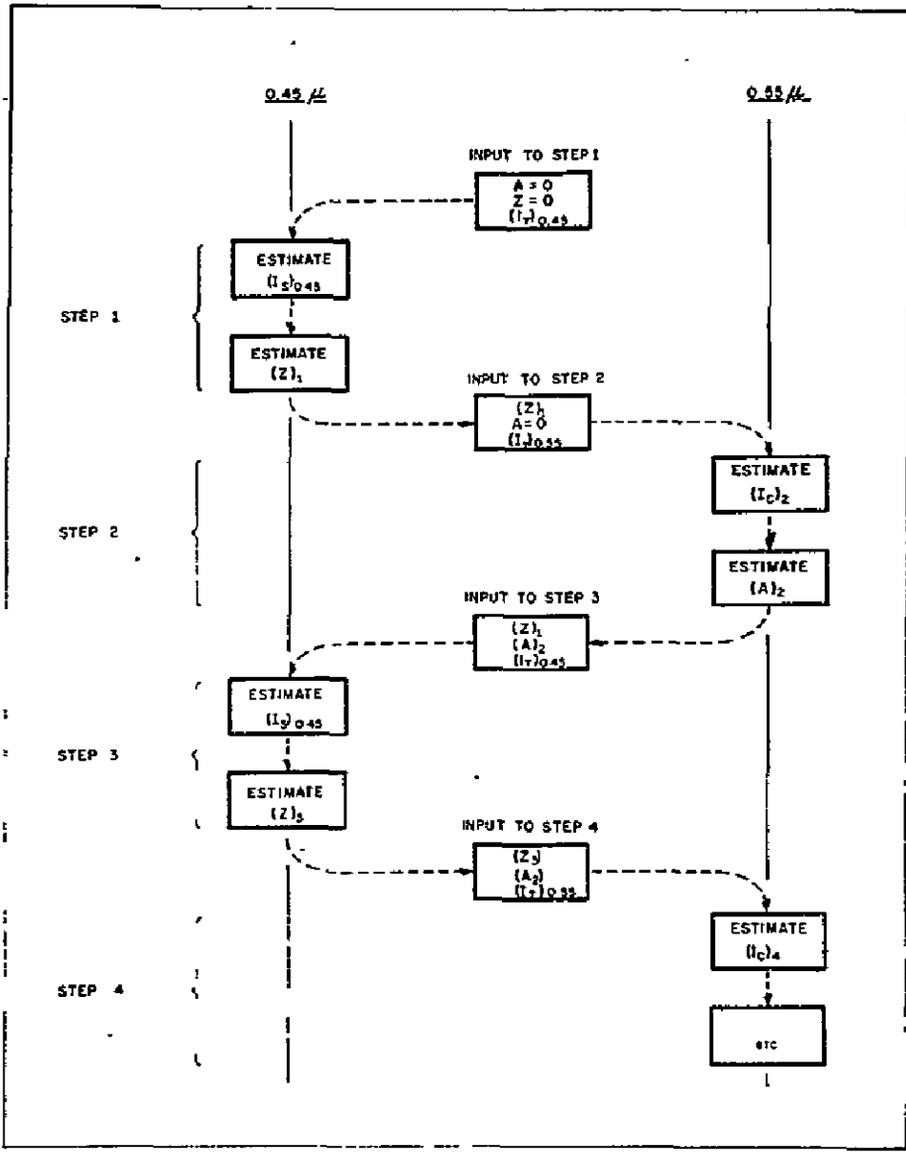


Fig. 15. Functional Diagram of the Radiance Inversion Technique.

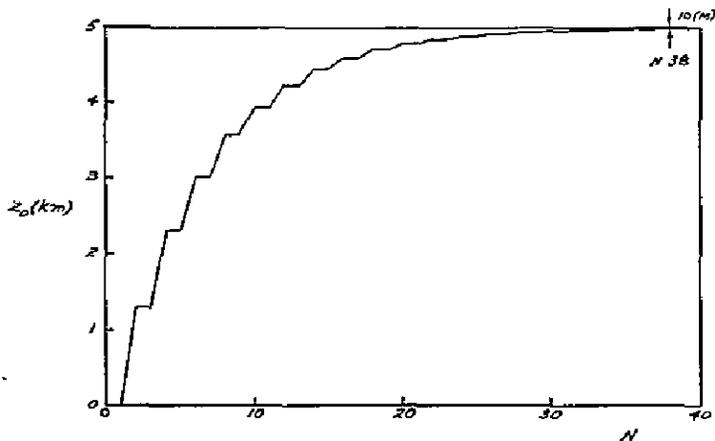


Fig. 16. Example of the Radiance Inversion Technique Showing Convergence on the Correct Height, $Z_0 = 5$ km.

that this technique converges rapidly for nearly all angular conditions of μ_0 and μ . For example, in the top left block of Fig. 17 is plotted the number of iterations required to converge within 10 meters of the correct height. It is seen in that figure that the number of iterations required varies only slightly with μ_0 and μ when the height is two kilometers and the reflectance is .05. That is, under these conditions the number of iterations varies from only 14 to 18. However, looking at the other blocks in Fig. 17, it is readily apparent that a greater number of iterations is required to converge under certain angular conditions.

b. Convergence Factors

The number of required iterations varies with angle. In Fig. 17 it is apparent that most rapid convergence occurs when the numerical value of μ_0 and μ are about equal. Similarly, slowest convergence occurs when either μ_0 is small and μ is large, or vice versa. This is particularly true with large surface reflectance values. In Fig. 9 we saw previously that the $\Delta I/\Delta Z_0$ values approached zero under these same angular conditions—which also explains the large number of iterations required for these particular angles (Fig. 17).

The number of iterations also appears to vary with height (the ordinate in Fig. 17). This is misleading because the iteration scheme begins at $Z = 0$. Therefore, we would expect convergence to the 2 km level more rapidly than the 10 or 15 km levels. Thus, the data blocks in Fig. 8 should not be compared vertically. There is, however, a natural reason why convergence within 10 meters of 15 km would be slower than within 10 meters of 2 km. It is that this system is sensitive to optical mass and we know that $\partial \rho/\partial Z$ is smaller than at 2 km. Thus, for a fixed ΔZ , a smaller resolution of optical mass is

Number of Iterations Required to Converge Within 10 Meters

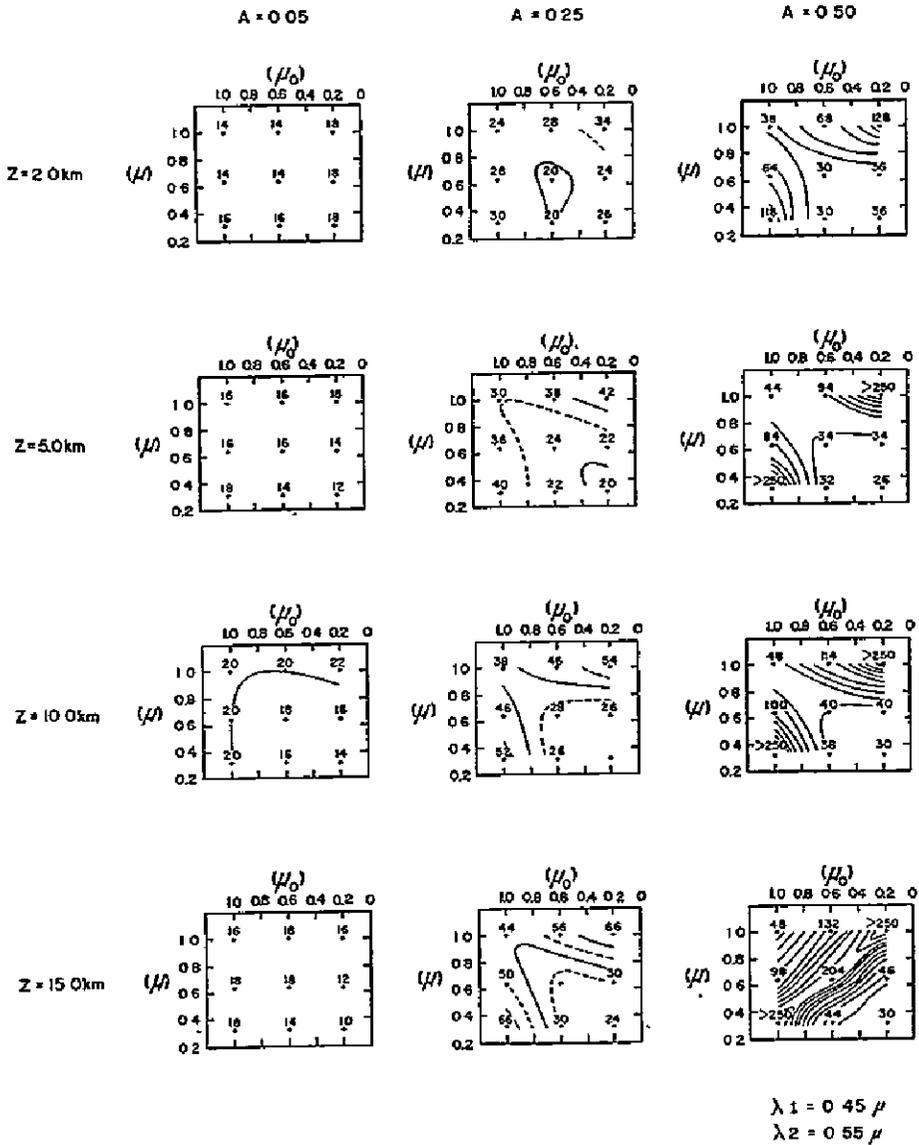


Fig. 17. Number of Iterations of Inversion Technique Required in Order to Converge within 10 Meters of the Correct Value, for Various Values of μ_0 , μ , A , and Z .

required at 15 km than at 2 km. This simply requires more iterations at 15 km.

Convergence can be obtained rapidly for nearly all angular conditions. Our results on the UNIVAC 1108 computer indicate an iteration rate of about 600 iterations per second. Thus, the complete inversion of two typical measurements requiring 20 iterations can be carried out in only 33 milliseconds.

IV. TESTS WITH SATELLITE DATA

At the present time the only available satellite bispectral radiance measurements for large portions of the earth are the ATS-III cloud camera data. These data are stored in three forms: (1) photographs, (2) digitized on magnetic tape, and (3) in analog form on magnetic tape. The photo in Fig. 18 is the ATS-III picture* for January 20, 1968 (1855 GMT) which has been used in this study to examine the observed brightness in the blue and green channels. The digital data were used in order to provide a quantitative measure of the brightness. The red channel had a major digitizing problem at this time and is not useful. Grid areas were selected for certain portions of the picture and are shown in Fig. 18. These areas have dimensions of 100 lines by 300 picture elements. (The entire picture is composed of 2400 lines and 8200 elements.) Near the subsatellite point these areas have dimensions of approximately 4.2° latitude on a side, or approximately 460 kilometers. Each area contains 30,000 individual ATS-III brightness samples.

In the photo (Fig. 18), it is possible to see cirrus moving out of the tropics from areas 16 through 18 toward higher latitudes. The ITC can be seen crossing areas 37 through 52. Cumuliform clouds associated with a large subtropical high pressure system can be seen in areas 61 through 86. An area of sunglint is also visible in the picture and is seen centered between areas 58 and 59.

For sea level pressure determination by inversion, it is necessary to remove the cloud noise from the 30,000 brightness measurements in each area. A discussion of how the noise was removed is presented in the next section.

During this work a digitizing problem with the data was noted. This problem occurred at the Rosman Ground Tracking Station. These digitizing errors were removed from the data and a discussion of the problem is given in the Appendix.

*This photo is a composite of the blue, green and red channels.

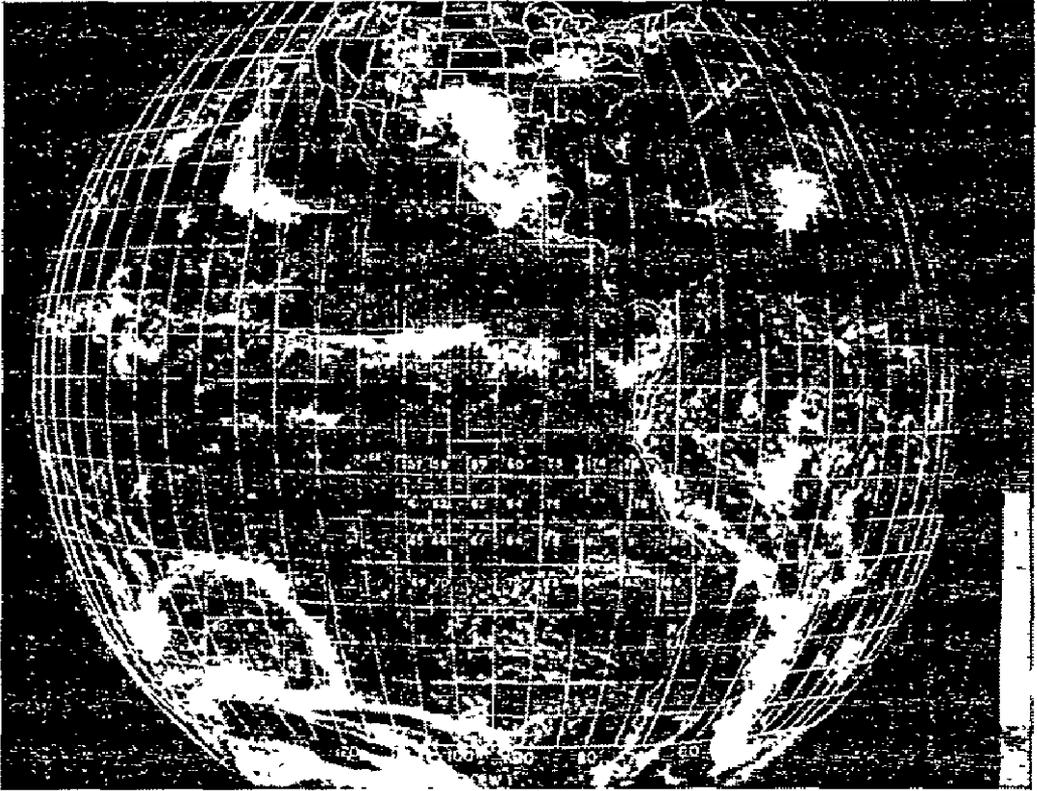


Fig. 18. ATS-III Spin-Scan Cloud Picture of January 20, 1968 (1855 GMT).

1. Removing Cloud Noise

Most satellite studies show that, given an arbitrary area of a few hundred square kilometers, there is a high probability of clouds occurring in the area. For example, studies by Smith (1969) (based on SIRS data) show there is partial cloudiness within an arbitrary 125 n. mile square at least 85% of the time. We also know from aircraft flights in the tropics that large cloudless areas are exceedingly rare. This is important to the present study because if we wish to discriminate surface brightness (for surface pressure) with a cloudless atmosphere, it is necessary to identify and eliminate the "cloud noise" data from our samples. This, of course, leaves us with fewer "clear" measurements than would be found in a cloud-free region of the same size.

A few of the 87 areas in Fig. 18 appear to be relatively cloudless. Two such areas are 10 and 48. The frequency of occurrence vs. brightness plots for these areas, shown in Fig. 19, indicate a near Gaussian distribution in brightness. This also suggests the area is relatively cloud-free. In other areas, however, as cloudiness increases, the right-hand side of the Gaussian

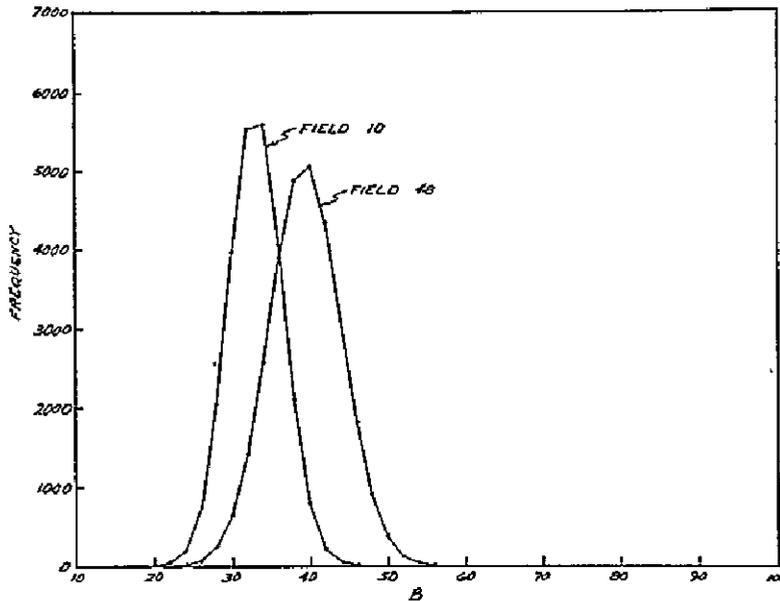


Fig. 19. Frequency of Occurrence vs. Observed Brightness with ATS-III
January 20, 1968 (1855 GMT).

distribution is skewed upward as shown in the top of Figs. 20 and 21—because of the increased frequency of occurrence of bright elements.

Based on these ideas, a method for attempting to remove cloud noise was determined as follows (and is shown in Figs. 20 and 21):

1. The frequency of occurrence vs. brightness data from ATS-III were corrected for the digitizing error, as described in the Appendix.
2. The data on the left-hand side of the Gaussian curve were used, and remaining data on the right-hand side (which includes cloud brightnesses) were ignored.
3. The data from the left-hand side of the curve were used to estimate the complete curve and its descriptive parameters including its mean brightness value. This was estimated by obtaining a least-square fit of the frequency of occurrence (F) vs. brightness (B) data to the expression,

$$F = \theta_1 \cdot e^{-\theta_2(B - \theta_3)^2} \quad (16)$$

θ_3 is the mean brightness.

These calculations were completed for each of the 87 areas in Fig. 18; the resulting mean brightnesses of the blue and green channels are shown in Figs. 22 and 23.

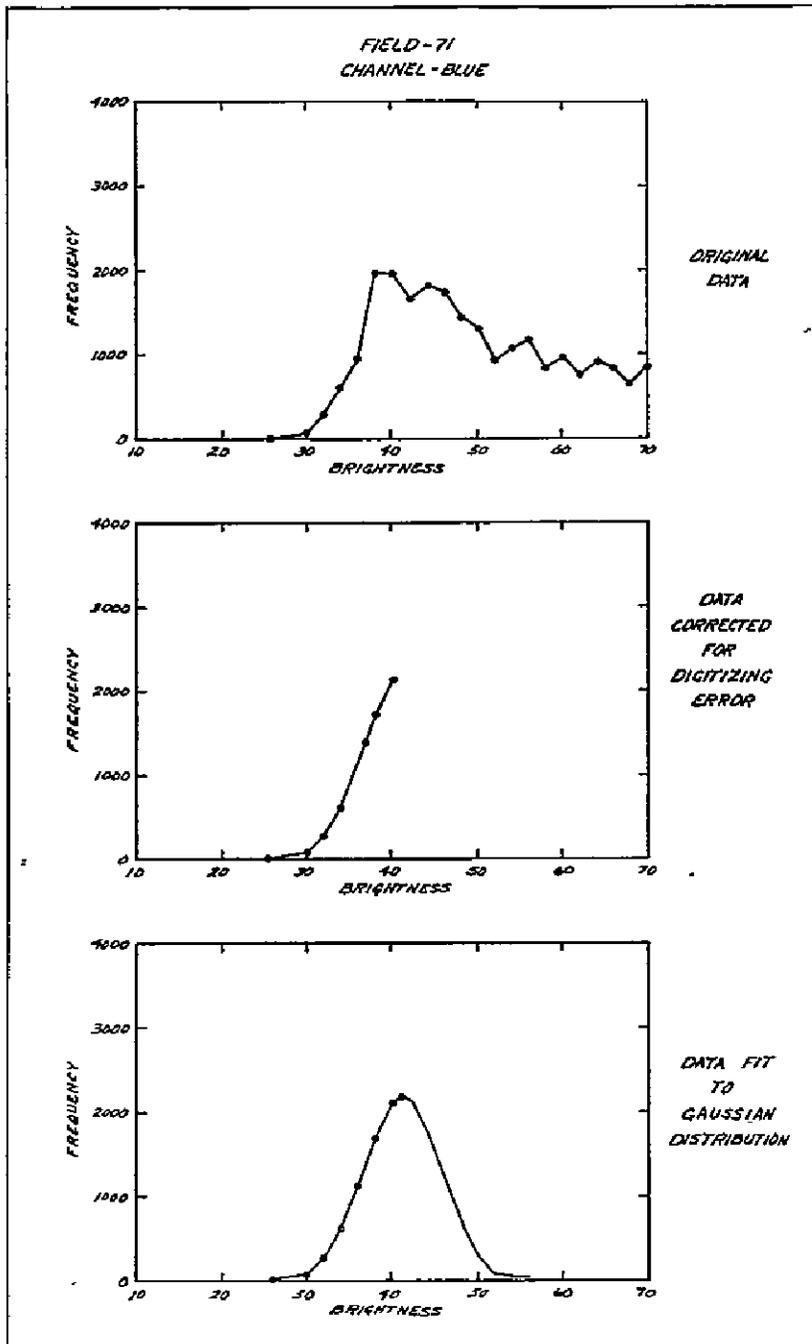


Fig. 20. Illustration of the Technique for Determining the Mean Brightness without Clouds, Field 71.

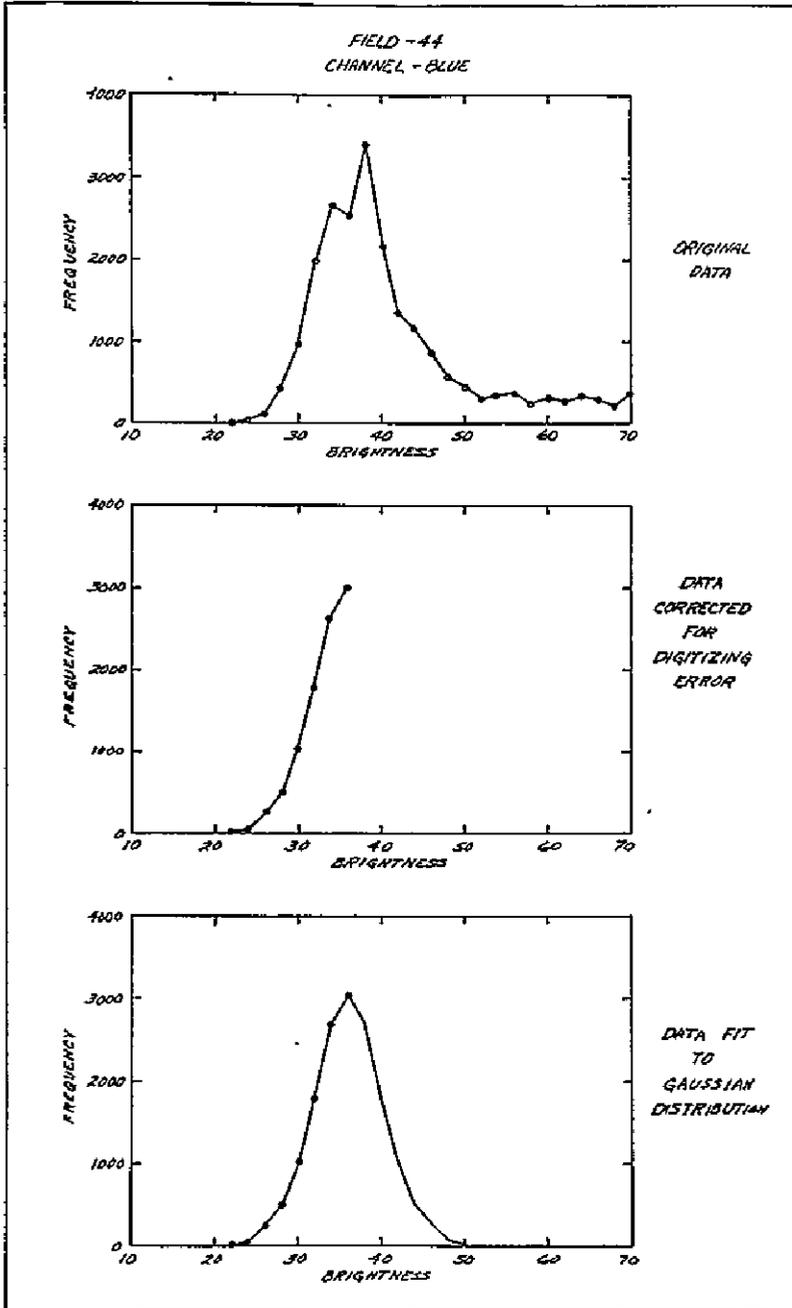


Fig. 21. Illustration of the Technique for Determining the Mean Brightness without Clouds.

If a horizontally homogeneous layer of clouds or aerosols extended across an area, such as in Fig. 18, we would find that the frequency of occurrence vs. brightness curve would be displaced to the right by an amount which depended on the brightness of the layer. Thus, the method described above for removing cloud noise would fail to remove the effect of a horizontally homogeneous (cloud) layer.

Similarly, small cloud elements, which are considerably smaller than the camera resolution and are uniformly distributed over the area, influence the left-hand side as well as other portions. The technique also fails to remove the effect of these clouds.

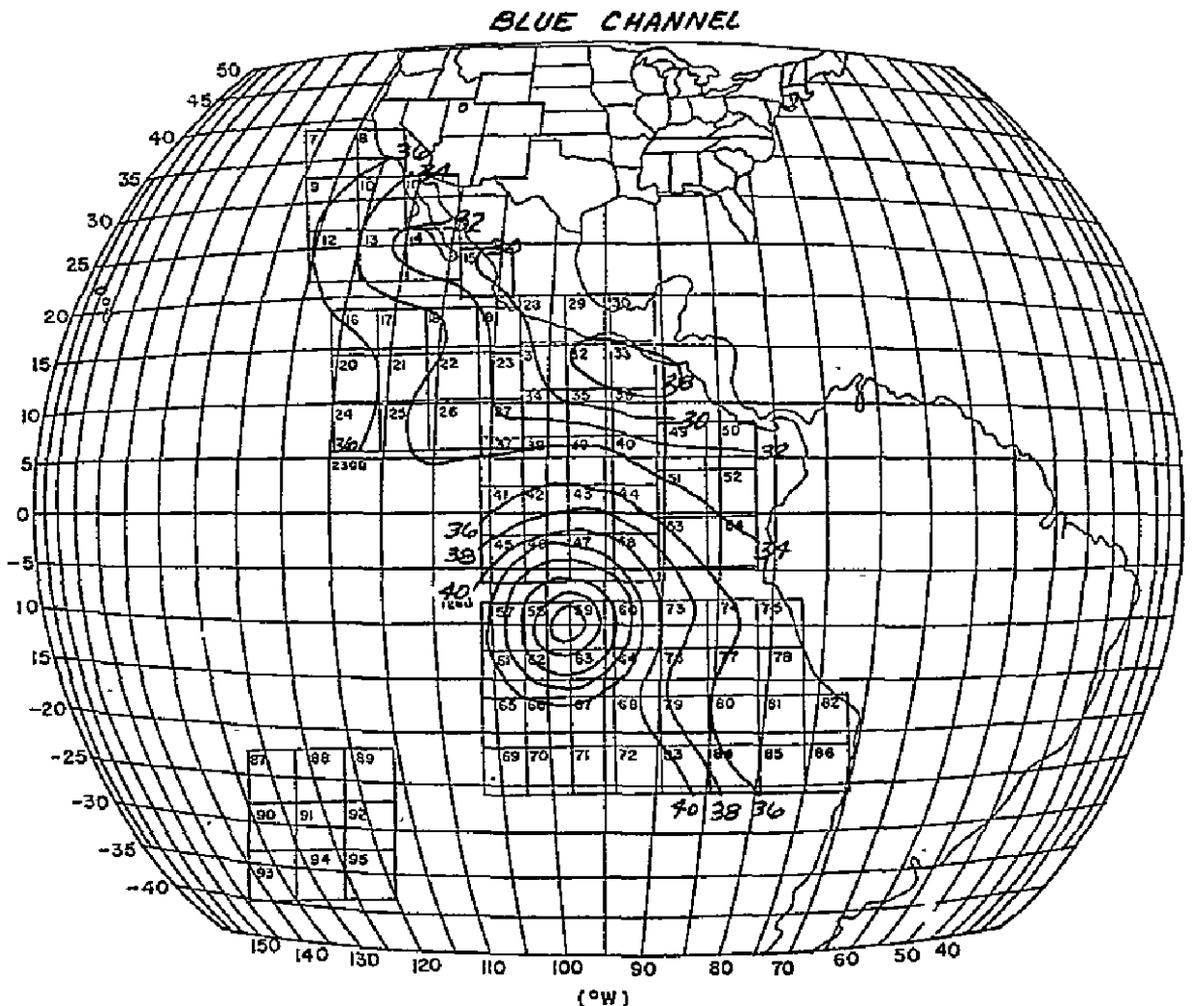


Fig. 22. ATS-III Observed Brightness without Clouds—Blue Channel.
January 20, 1968 (1855 GMT).

There is evidence in Figs. 22 and 23 that these two cloud situations may be influencing our calculation of "cloudless" conditions with ATS-III data. In areas 16-21, for example, where there appears to be considerable cirrus, the brightness values for both channels are larger than would be expected (Figs. 22 and 23). Similarly, in areas 37 - 40, which includes the ITC, the brightness values show considerable scatter (Table 3). The brightness in areas 7 and 9 also appears to be slightly greater than would be expected and may be influenced by clouds. Thus, the present technique (used with ATS digital data) appears to be limited in completely removing the effect of clouds.

Some attempt has been made to determine a more effective means (than the technique outlined above) for removing the influence of cloudiness. One would imagine, for example, that perhaps a portion of the left-hand side of the Gauss-

GREEN CHANNEL

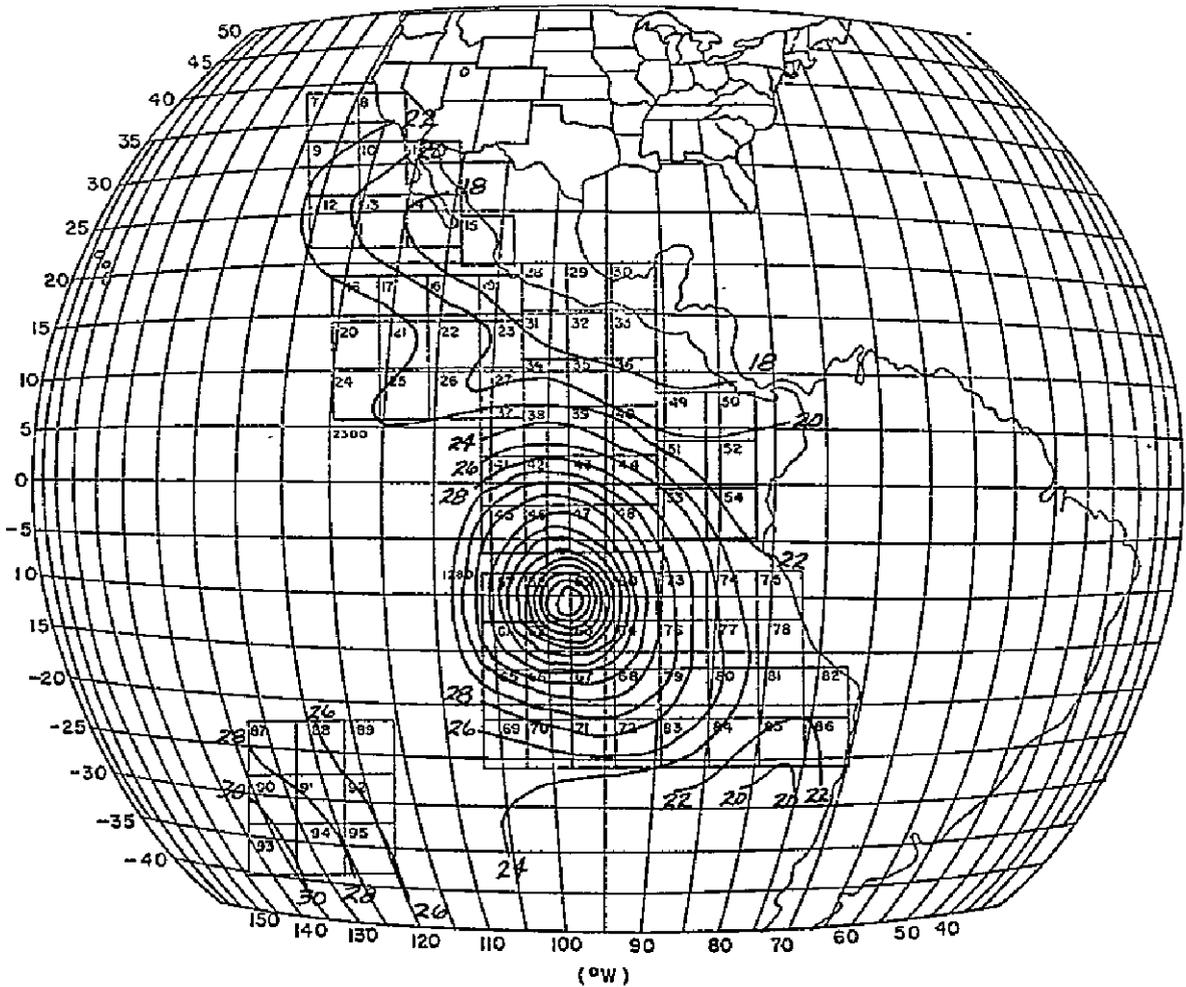


Fig. 23. ATS-III Observed Brightness Without Clouds—Green Channel. January 20, 1968 (1855 GMT).

Table 3

Field No.	Mean Brightness		Field No.	Mean Brightness	
	Blue Channel	Green Channel		Blue Channel	Green Channel
7	37.93	27.15	50	31.62	19.51
8	36.47	27.49	51	35.38	26.42
9	36.71	23.38	52	33.11	22.60
10	33.38	4.00	53	16.00	25.48
11	33.56	19.85	54	31.60	19.83
12	35.01	20.95	57	43.96	38.02
13	31.58	18.56	58	49.95	51.94
14	31.79	17.42	59	49.57	51.54
15	29.76	17.49	60	43.56	37.21
16	36.05	24.88	61	42.89	35.55
17	34.06	21.21	62	48.96	43.82
18	34.44	20.02	63	46.38	43.42
19	31.10	17.79	64	41.72	34.85
20	39.80	26.14	65	41.95	29.50
21	34.85	23.96	66	41.35	30.48
22	32.65	21.09	67	42.49	32.72
23	32.05	20.21	68	41.24	31.88
24	37.55	28.15	69	40.56	25.10
25	34.07	21.12	70	40.22	25.89
26	32.21	19.91	71	41.09	27.16
27	32.05	20.51	72	45.60	29.08
28	29.48	16.99	73	38.73	29.63
29	29.39	17.48	74	36.26	26.64
30	31.05	17.01	75	35.85	26.03
31	28.71	16.67	76	37.08	27.90
32	27.35	16.84	77	35.71	26.92
33	27.09	21.38	78	37.55	25.04
34	31.98	20.51	79	42.76	28.81
35	30.55	19.16	80	35.51	29.17
36	29.18	18.03	81	35.64	32.65
37	35.60	10.00	82	36.41	23.24
38	36.94	31.67	83	40.33	25.83
39	24.00	12.00	84	36.23	21.78
40	35.73	23.44	85	34.27	20.91
41	35.39	28.03	86	36.54	23.46
42	38.21	31.29	87	41.33	23.78
43	41.02	34.54	88	44.03	26.24
44	36.08	29.38	89	43.09	27.34
45	40.04	33.35	90	45.12	32.37
46	43.42	38.03	91	43.21	28.00
47	42.98	36.71	92	42.40	26.68
48	39.68	33.27	94	51.86	27.41
49	31.81	18.59	95	43.12	32.36

ian curve would be a better predictor of the mean brightness (θ_3) than is the entire curve. We have made two separate studies on this problem.

The first is that we have looked for a "signature" in the Gaussian curve which might be descriptive of the effect of clouds. This has been done by plotting the frequency vs. brightness data on a probability scale. Such a plot should be linear if the data have a normal distribution. Shown in Fig. 24 are the plots for areas 10, 32, 36 and 48 which appear in Fig. 18 to be relatively cloudless. These curves should be compared with the curves representing cloudy areas (both stratocumulus and deep tropical convection) which are shown in Figs. 25 and 26. In all areas the curves are quite linear over most of the range, except near the highest frequency values where they become nonlinear--particularly for the "clear" areas (Fig. 24). It should be pointed out that these curves have been obtained by normalizing the observed frequency values on the left side of the Gaussian curve to the maximum frequency value, whose normalized value is 0.50. The maximum value may or may not correspond to the true peak of the Gaussian curve depending on the effect of clouds or any other nonrandom factor. To normalize the data properly, however, the maximum value and the peak of the Gaussian curve should correspond exactly. Therefore, the position of a curve on the graph depends on the single value on which it has been normalized. The importance of all this is that, as a result of the normalization, it is not possible to linearly extrapolate the curve to $F = 0.50$ in order to determine the mean brightness (θ_3). It appears, from this study, that the maximum frequency value which we observe may occur at a slightly greater brightness than the brightness at the true peak of the Gaussian curve. This is most likely an effect of cloudiness.

The question which this suggests is, can we improve in our estimate of the mean brightness (θ_3) if we take successively fewer data points along the left side of the curve? We have evaluated this question by calculating θ_3 in the usual amount by using all data from the smallest to the maximum frequency value, or a total of N observations. These calculations were repeated for $N - 1$, $N - 2$, etc. observations. The results are shown in Table 4. We have made the calculations on four areas which appear to be relatively cloud-free. Compared to the mean brightnesses with N observations, there appears to be no systematic change in the brightnesses with $N - 1$, $N - 2$ and $N - 3$ observations. However, for $N - 4$ observations the mean brightness values are systematically less. This may be due to the small number of data points at $N - 4$, because at $N - 5$ there were so few data points that the least-squares iteration would not converge in determining θ values.

Thus, with the present ATS-III data, which has a limited number of data points to define the curve, the best definition of the mean brightness without clouds is obtained by using all data points on the left side of the curve. These data suggest, however, that when other data are available with greater brightness resolution, it may be advantageous to disregard data points near the peak of the curve in order to better determine the mean brightness (θ_3).

Table 4

Area No.	N-5	N-4	N-3	N-2	N-1	N
	Mean Brightness (θ_3)					
8	-	29.6	39.2	36.7	36.1	36.5
10	-	28.8	32.2	36.6	33.3	33.4
13	-	28.6	29.9	30.7	35.0	31.6
19	-	27.3	30.3	29.9	28.7	31.1

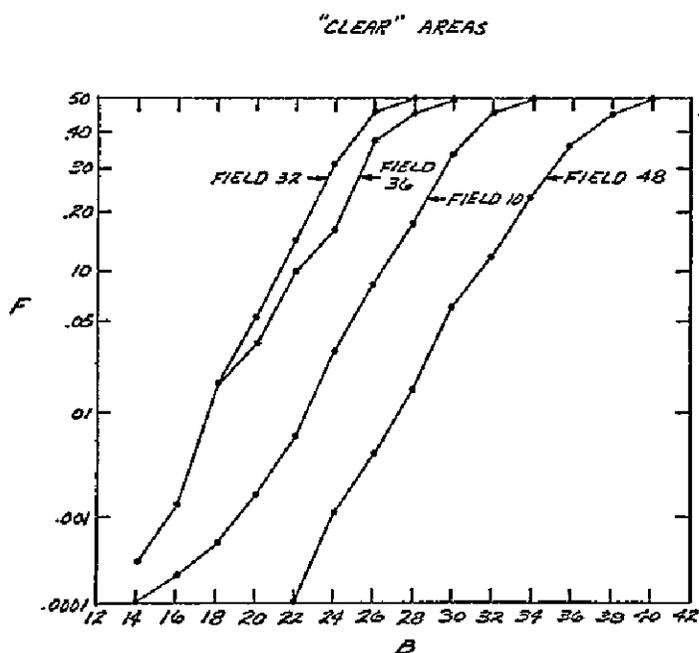


Fig. 24. Gaussian-Scale Plots of the Frequency of Occurrence vs. Brightness from the Lowest Brightness to the Brightness at N for the Relatively Cloud-free Areas 10 and 48.

CLOUDY AREAS
(STRATOCUMULUS)

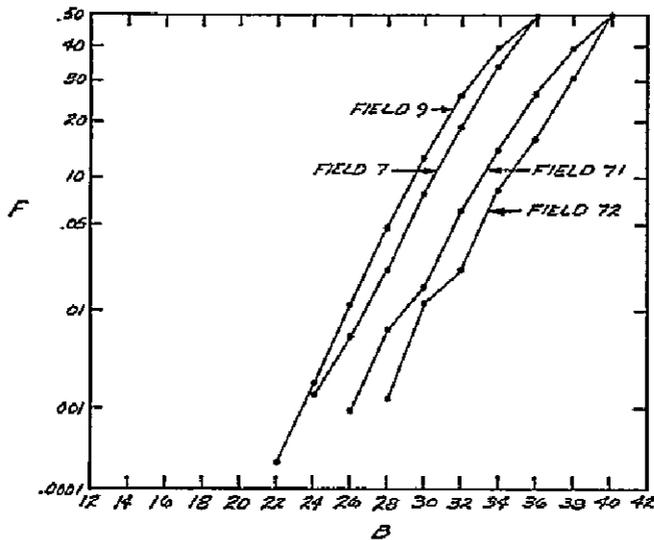


Fig. 25. Gaussian-Scale Plots of the Frequency of Occurrence vs. Brightness from the Lowest Brightness to the Brightness at N for Areas 7, 9, 71 and 72 which are partially covered with stratocumulus clouds.

2. Observed Radiance Distribution and Comparison with Theoretical Radiance Values

The mean brightness values (θ_3) from the previous section have been determined for each area (7-95) in Fig. 18 by the method described above for removing cloud noise. These mean brightness values have been plotted and isophotes are drawn in Figs. 22 and 23 for the ATS-III blue and green channels, respectively. In both figures, we see a strong sunglint centered between areas 58 and 59. The brightness gradient in the sunglint is greater in the green channel than in the blue. Apparently this is because there is less attenuation of the sunglint in the green channel.

In order to compare the ATS-III observed values with theoretical radiance values, we have calculated μ_0 , μ and ϕ for January 20, 1968 (1855 GMT). The values are shown in Fig. 27. From these parameters we have calculated the upward radiance for the blue channel, assuming $\lambda = 0.435$ microns, and a reflectance of 0.10 at the lower boundary (Krinov, 1947). Dave and Warten's computer program (1968) was used for these calculations. The result is shown in Fig. 28. We see limb brightening over most of the earth's disc, except decreasing radiance toward the northeast—due to the large zenith angle of the

sun in that area. This figure of the theoretical radiance should be compared with Fig. 22 which is the ATS-III observed radiance (blue channel). This particular time was selected for comparison in order to minimize the errors due to aerosols which are present in the observed data and not in the model.

To facilitate a comparison of Figs. 22 and 28, a brightness cross section has been taken along the 100th meridian and is shown in Fig. 29.

Data for both the blue and green channels have been included. The most striking feature is that sunglint shows up markedly in the observed data but is absent in the theoretical calculations, of course, because the model assumes a diffusely reflecting lower boundary. The observed brightness values for area 38 (at 4.5°N) appear to be larger than expected due to the ITC cloudiness which, apparently, is not fully removed from the data.

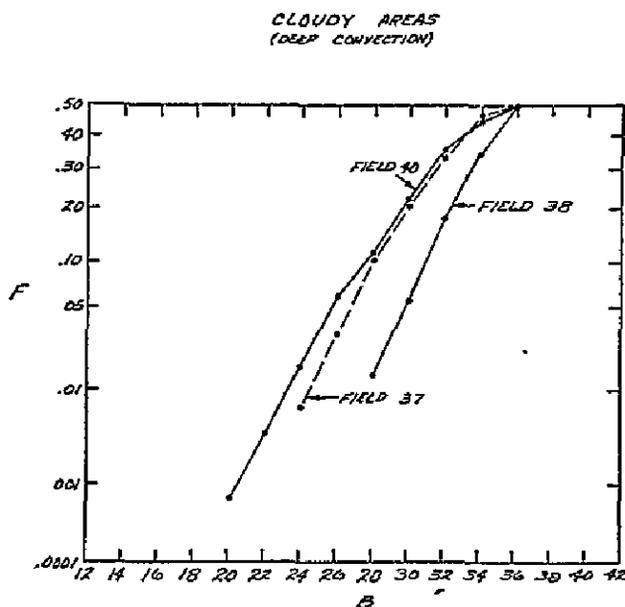


Fig. 26. Gaussian-Scale Plots of the Frequency of Occurrence vs. Brightness from the Lowest Brightness to the Brightness at N for Areas 37 and 38 which are heavily covered with deep convection.

2

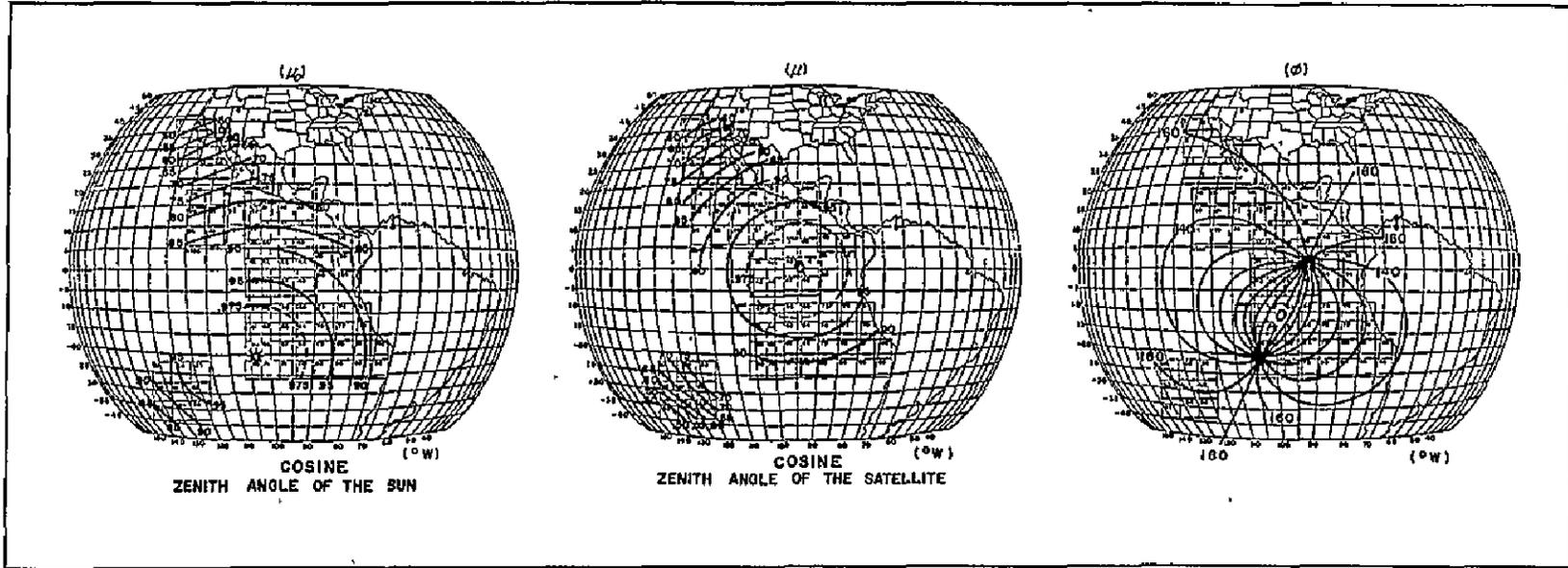
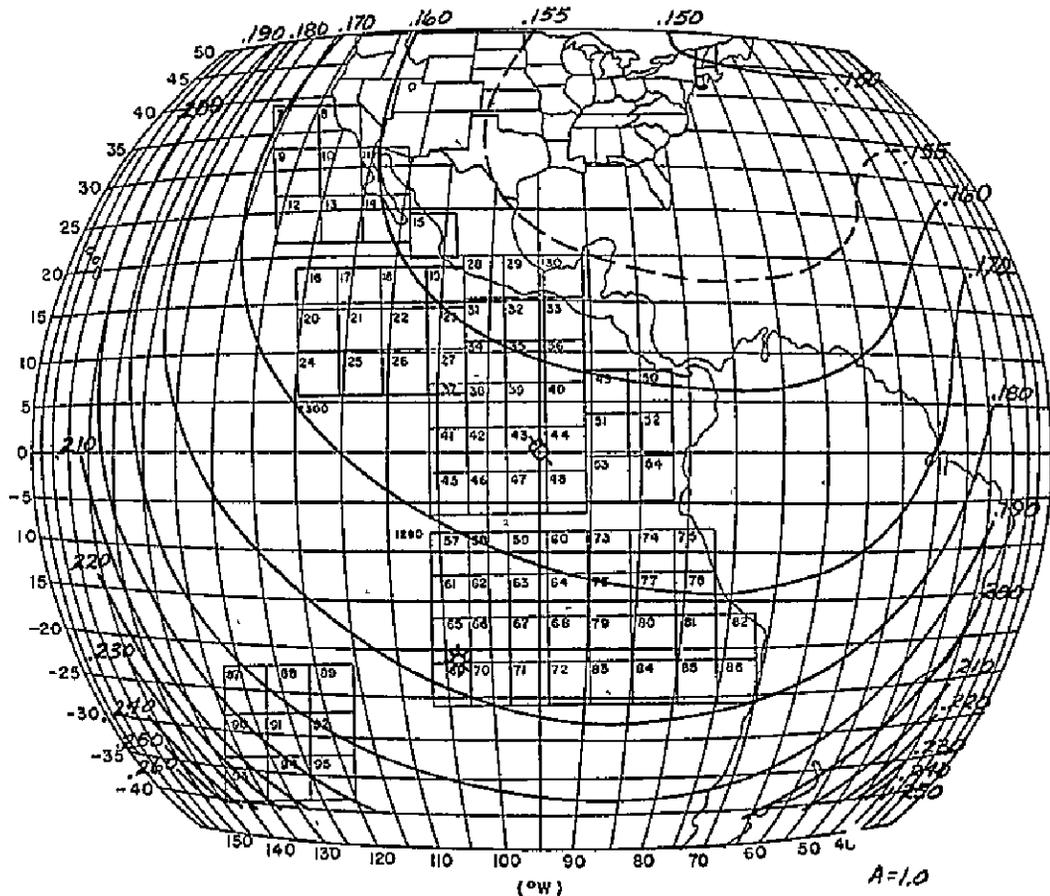


Fig. 27. Angular Values for Points on the Earth's Disc on January 28, 1968 (1855 GMT). Sun (-20°S , 104°W). Satellite (0°N , 94.9°W , altitude above Earth's Center is 42,139 km.).

THEORETICAL RADIANCE
 DAY 20, 1968 (1855 GMT)



56

Fig. 28. Theoretical Radiance from a Rayleigh Atmosphere and Diffusely Reflecting Surface ($A = .10$) on January 28, 1968 (1855 GMT). Sun (-20.2° , 104°W). Satellite (0°N , 94.9°W , Altitude above Earth's Center is 42,139 km.).

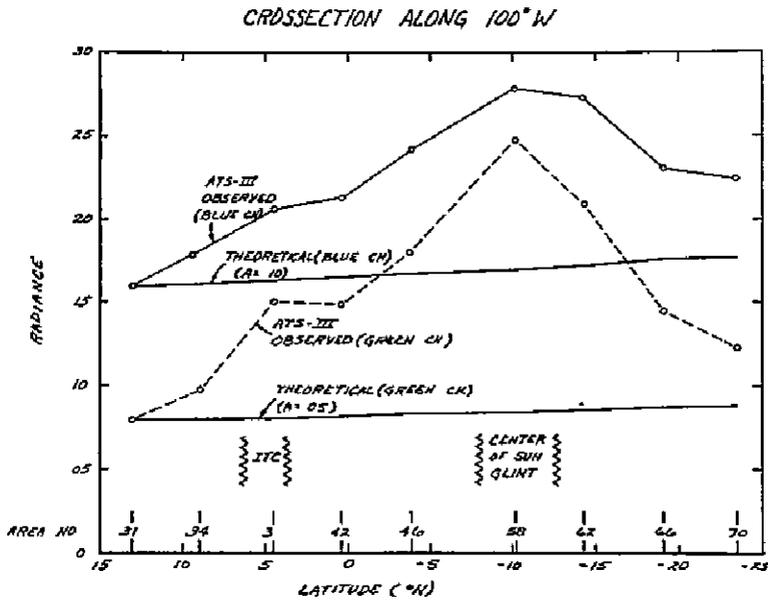


Fig. 29. Comparison of Theoretical and Observed Radiance along 100°W, January 20, 1968 (1855 GMT).

3. Test of Inversion Technique on ATS-III Data

In order to test the inversion techniques on real data, we have used the ATS-III digital data for January 20, 1968 (1855 GMT). We have attempted to remove the cloud noise from these data, as discussed in the previous section.

The observed blue and green channel brightness values (Table 3) were normalized to the theoretical radiance values for area 31. The latter values are:

$$I_{TB} = .16041$$

$$I_{TB} = .07950$$

and are based on the following assumed values

Blue Channel

A = .10
 τ = .250
 λ = .435 (microns)
 Pressure = 1013.3 (mb)

Green Channel

A = .05
 τ = .110
 λ = .535 (microns)
 Pressure = 1013.3 mb.

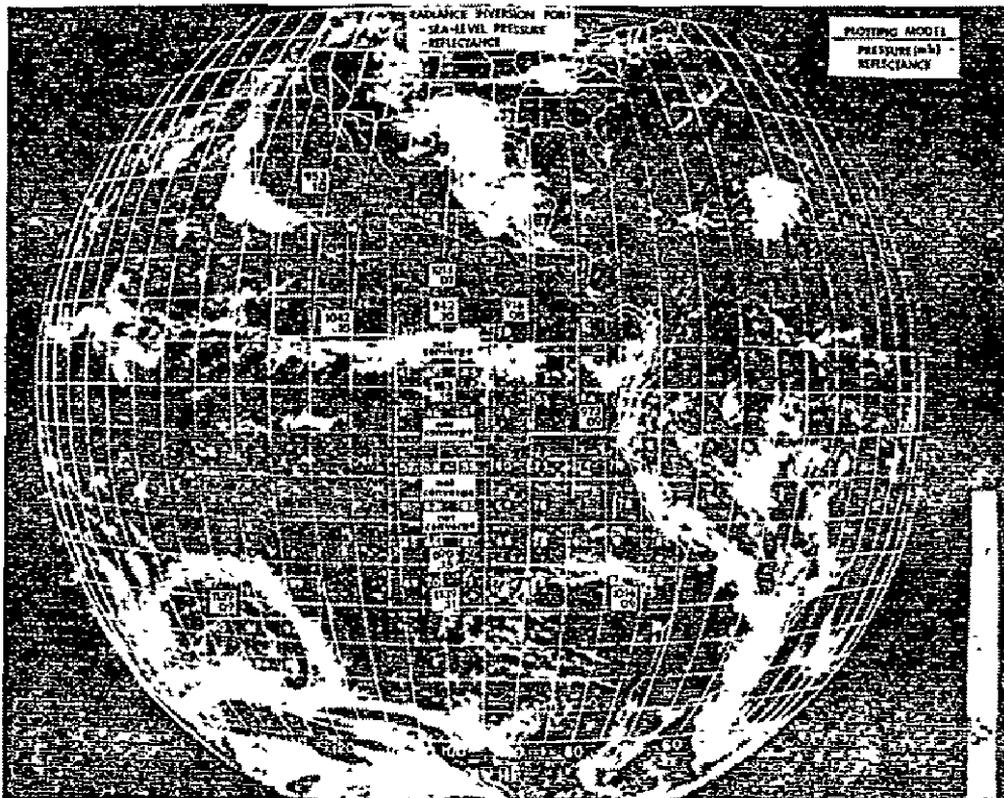


Fig. 30. Result of Inversion of ATS-III Data (Blue and Green Channels) to Obtain Sea Level Pressure and Reflectance.

The resulting normalized ATS-III data are tabulated in Table 5. These data were then used to test the inversion technique. A total of 15 areas were used in this test. We have shown the areas in Fig. 30.

A wide range of locations and brightness conditions was included in the test. As shown in Fig. 30, in the three areas along 100°W which are most affected by sunglint, we did not obtain convergence in the inversion. Obviously, this is because the model is for diffuse reflectance. Also, convergence could not be obtained for area 42, which is nearly covered by cloudiness of the ITC. However, for the remaining 10 (independent) areas, convergence was obtained rapidly for each case. Results of the inversion shown in Fig. 30 are also summarized in Table 5. If we disregard the two areas (42 and 46) which have largest errors (perhaps also due to sunglint), the variance of sea level pressure obtained by inversion is 80 mb.

The reflectance values obtained by inversion are quite constant at .07 - .10 away from the sunglint. Areas 42 and 66 show that somewhat larger reflectance values (.18 and .15) are found along the edges of the sunglint.

Table 5

Area No.	Lat. (°N)	Long. (°W)	Observed Radiance		Inversion Results		Observed Sea Level Pressure
			Blue Channel	Green Channel	Reflectance	Sea Level Pressure (mb)	
13	24.0	117.0	.14636	.10181	.095	953	1021
31	13.0	100.6	.16041	.07950	.070	1013	1013
25	8.0	112.6	.15790	.11589	.096	1042	(1013)
34	8.9	100.4	.14819	.11254	.098	942	(1013)
36	8.9	92.5	.13521	.09895	.083	916	1012
54	-2.5	83.7	.14645	.10882	.090	973	1012
70	-23.5	100.2	.18638	.14208	.108	1139	1024
85	-22.7	77.5	.15879	.11473	.089	1014	1016
87	-25.0	130.0	.19151	.13048	.090	1139	1012

Variance of Pressure = 80 mb.

There does not appear to be systematic errors in the inversion result as a function of distance from the normalizing area (in this case area 31). This is supported by the fact that reasonably good inversion results are obtained in areas 85 and 87 which are 2500 to 3000 miles away from area 31. Pressures of 1014 and 1139 mb and reflectances of .09 were obtained in these areas. As indicated in Table 4, these pressures are in error by 2 and 127 mb, respectively.

V. POSSIBILITIES FOR IMPROVED PRESSURE AND HEIGHT RESOLUTION

There are a few techniques which could improve our ability to discriminate the pressure or height of the lower boundary. These will be discussed individually.

1. Higher Digitizing Resolution

The signal to noise ratio (S/N) of the PM tube is fixed by the input radiance—and therefore there is no capability of improving this basic S/N value. However, the present ATS-III is being digitized from 0 - 255 for the full brightness range of the camera, and we have seen in Fig. 22 that for a cloudless atmosphere the brightness rarely exceeds 40 on the scale from 0 - 255. Thus, the ground station gain could be increased by a factor of 5.

In addition, the ATS-III A/D converter is an 8 bit system (of which our analysis has used only 7 bits—the lowest bit was unusable). It is possible to use instead a 9 or 10 bit converter, and thereby raise the resolution by a factor of 2 or 3 over that used in the present analysis.

Thus, by including both a higher gain and larger bit converter, the resolution could be improved by a factor of 10 to 15. This means, for example, that instead of having 8 points to resolve the ocean brightness (Gaussian) curve (shown in Fig. 21) we would have from 80 to 120 points.

2. Spectral Separation for Radiance Measurements

The work presented in Chapter III on error estimates for cloud height and sea level pressure were based on an assumed radiance value at 0.45 microns, if, instead, a shorter wavelength was used, the $\Delta I/\Delta Z_0$ would increase due to Rayleigh's λ^{-4} law.

If, for example, the measurement was made at 0.36 microns, Table 2 shows that the radiance sensitivity with height ($\Delta I/\Delta Z_0$) would increase by a factor of about 2 compared to that at 0.45 microns. Figure 31 (which should be compared with Fig. 8) shows this improvement.

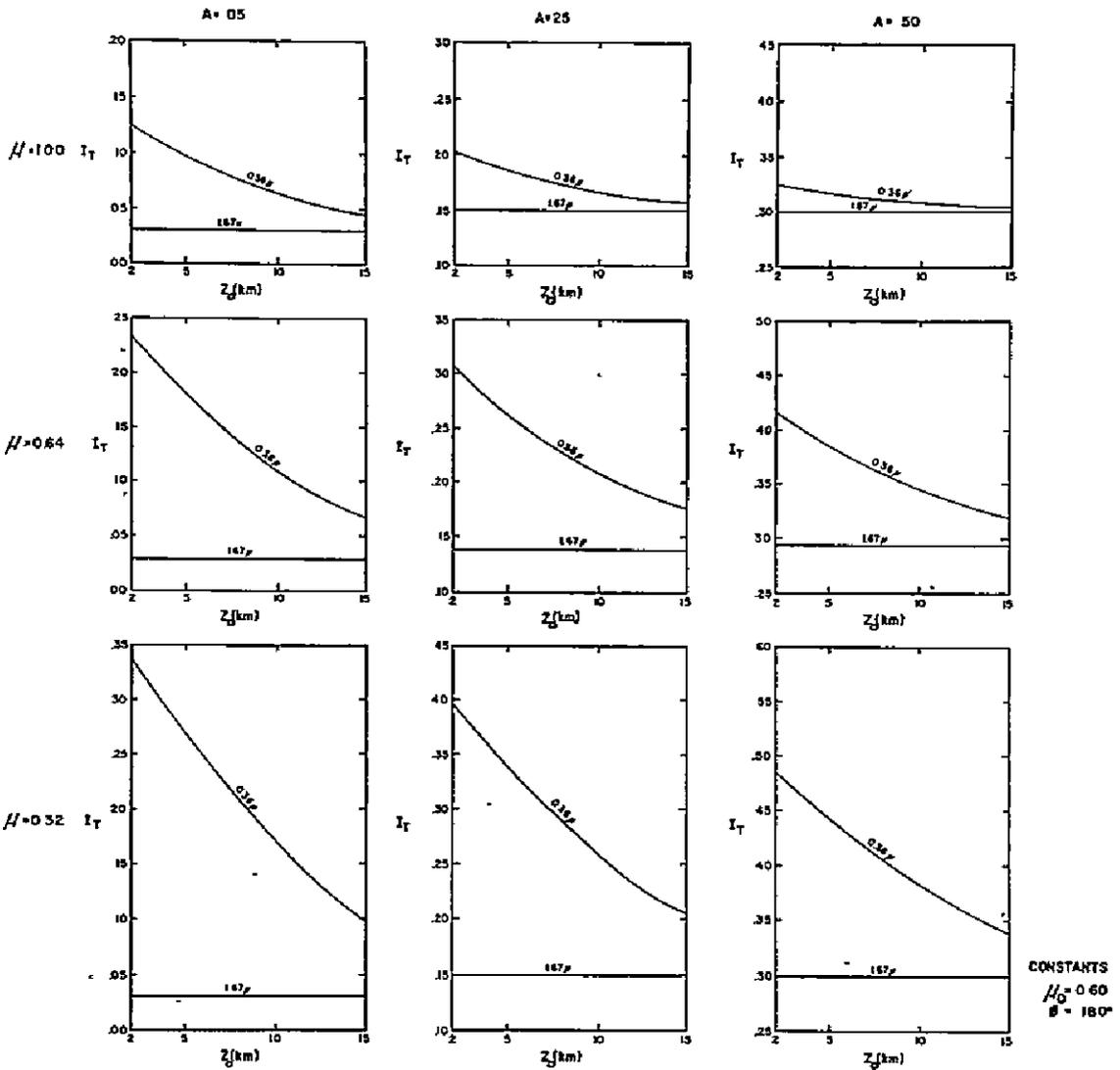


Fig. 31. Radiance Dependence on Z_0 , for Certain Constant Values of μ_0 , μ , ϕ , A and Wavelength Values of 0.36 and 1.67.

The longer wavelength measurement is used to determine surface reflectance. Thus, it should be independent of atmospheric scattering. If 1.67 microns were used, Fig. 31 and Table 2 show it is virtually independent of Rayleigh scattering.

Obviously, the effect of aerosol on the radiance is very large at this wavelength (1.67 microns) compared to the effect of molecular scattering. For example, Elterman (1965) tabulations of the Rayleigh optical thickness (τ_R) and total optical thickness (τ_T) (including aerosols and O_3) show that τ_T is greater than τ_R by nearly 2 orders of magnitude. This is because τ_R is extremely small. The important question is, however, how large is the radiance from aerosols compared to the radiance from the surface, because it is the reflectance of the surface which must be determined in the near infrared.

Perhaps two measurements will ultimately be required here—one to detect aerosols and one to determine the surface reflectance.

3. Large Samples to Reduce Electrical Noise

It was previously pointed out (Chapter III) that the error in a single measurement can be improved greatly by averaging a large number of measurements in a grid area to reduce PM noise. Since this noise is random, it reduces by the factor $1/\sqrt{n}$, where n is the number of measurements.

If a pressure measurement is obtained over an area 460×460 km, the pressure error is reduced by a factor of 100. If a cloud measures 30×30 km, the error in cloud top height (Z_0) is reduced by a factor of 17.

4. Polarization Measurements

Electromagnetic radiation is vectorial and requires four parameters to fully describe its state. One of these parameters is intensity (I) and another is the degree of polarization (P). The latter is defined in terms of Stokes parameters (Sekera, 1961) as,

$$P = \frac{[(I_r - I_e)^2 + U^2]^{1/2}}{I_r + I_e} \quad (17)$$

where I_r and I_e denote the intensity components perpendicular and parallel respectively to the vertical plane that contains both the local vertical and the direction of observation, and U represents the third Stokes parameter which determines the orientation of the plane of polarization.

When the direction of maximum intensity coincides with the direction of either I_r or I_e , then a simple case exists in which P is described by only the maximum intensity (I_{\max}) and minimum intensity (I_{\min}).

$$P = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} . \quad (18)$$

Certain cases of atmospheric polarization observed at satellite altitude are easily visualized. For example, if the surface reflectance is zero and if only single scattering occurs, then we would expect complete polarization at a scattering angle of 90 degrees and no polarization at zero degrees. A surface reflectance (unpolarized) of greater than zero will reduce the polarization observed at satellite altitude, compared to the polarization of a zero reflectance surface.

For most cases of satellite observation of polarization, it is not easy to visualize how the angular and reflectance factors affect the polarization. Therefore, we have calculated the polarization from the atmosphere, as observed by a synchronous satellite. (Similar, more complete cases of P have been made by Sekera, 1961, for the satellite at an infinite distance from the planet.)

Shown in Figs. 32 and 33 is the polarization from Rayleigh atmosphere with an unpolarized surface reflectance of 0.10. The optical thickness of 0.2238 was selected because it corresponds to the Rayleigh optical thickness of the earth's atmosphere at a wavelength of 0.45 microns.

The most important difference between Figs. 32 and 33 is the degree of polarization. Note that when the scattering angle is near 90° (Fig. 32), the polarization is large; in this case it is between 0.50 and 0.80 over the earth's disc. However, when the scattering angle is near zero (Fig. 33), the polarization is small—ranging from zero to 0.05 over the earth's disc. This is the important property of polarization. It also accounts for the fact that P is largest near the terminator at 0600L time at the subsatellite point. There is also an increase in polarization toward the earth's limb when A is small (such as in Fig. 33). This increase is even larger with larger values of A .

In Fig. 12, we saw that near the terminator it was difficult to determine sea level pressure by the inversion of radiance alone. We have seen in Fig. 32 that polarization is largest near the terminator. Thus, there is an excellent opportunity to use the inversion of polarization in the region near the terminator, where radiance inversion is poor.

A limitation on the above is that a water surface is a polarizing reflector. Therefore, if polarization is to be used effectively in an inversion scheme, there must be a method of separating the polarization from the atmosphere and that from the surface.

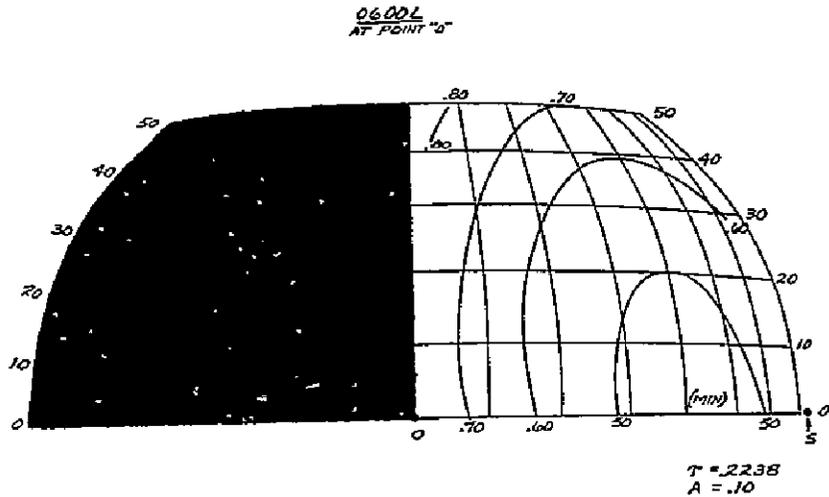


Fig. 32. Polarization for a Rayleigh Atmosphere and Nonpolarization Lower Boundary with $A = .10$ and $\tau = .2238$ and 0600 Local Time at the Sub satellite Point. Satellite Height Above the Earth's Center is 42,139 km

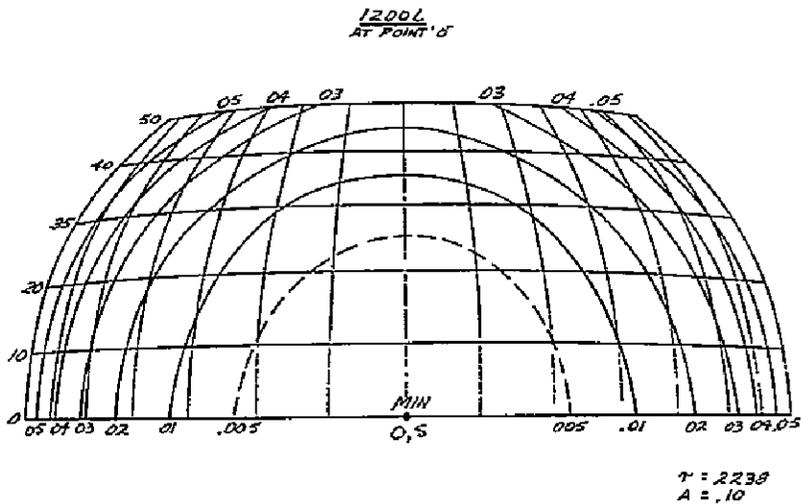


Fig. 33. Polarization for a Rayleigh Atmosphere and Nonpolarization Lower Boundary with $A = .10$ and $\tau = .2238$ and 1200 Local Time at the Subsatellite Point, Satellite Height Above the Earth's Center is 42,139 km.

VI. SUMMARY AND RECOMMENDATIONS

This study coincides with a general effort by many scientists to explore the feasibility of obtaining measurements to describe the state of the earth's atmosphere by remote sensing. In order to remotely sense any atmospheric parameter requires (1) a suitable radiance inversion scheme, either analytical or iterative, for the model atmosphere or conditions assumed, and (2) a set of radiance data which are representative of this model atmosphere or conditions assumed.

The work that is presented here considers the possibility of obtaining sea level pressure and "cloud" top height for a molecular atmosphere without aerosols and absorbing gases and for a lower boundary of the atmosphere which is a diffuse reflector. The study shows that such measurements are feasible for this model atmosphere and surface. Unfortunately, however, such a model atmosphere is only a first approximation to the earth's atmosphere. Eventually studies similar to this must be carried out with model atmospheres which are more accurately representative of the earth's atmosphere if we are to use earth atmosphere data for an inversion. These studies cannot be carried out now because the theoretical description of multiple scattering for these complex atmospheres has not been completed.

In order to test the inversion of radiance for the simple model atmosphere and surface, which we have assumed, we have used the blue and green channels of the ATS-III SSCC data. Obviously, this data does not match the model. In addition we have noted a digitizing problem in the ATS-III digital data and a relatively low camera gain setting which does not allow suitable resolution of the low (ocean) brightness values which are used in the inversion.

Inasmuch as these data are the only data available on which to test the radiance inversion, we have used these data in an iterative inversion technique to obtain sea level pressure and reflectance. The results of this test are encouraging in that we are able to obtain a solution over most of the earth's disc (viewed from synchronous altitude) except in the sunglint from the ocean. This region of sunglint is about 15 - 20 degrees latitude in diameter. The interesting point is that this is precisely the area where we should not obtain a solution, because the sunglint reflectance is specular. At other locations, the inversion is obtained satisfactorily; however, the variance of the calculated vs. observed sea level pressure is 80 millibars. This large error is due in part to the problems with the ATS-III data and in part due to the incompatibility of the model and the data. It is not possible to distinguish the individual contribution of each of these two error sources.

This problem will be resolved when theoretical studies are able to include the effects of aerosols and specular reflectance on the multiple scattering from

the atmosphere, and when improved satellite data with higher brightness resolution (at low ranges) and better digitization are available.

APPENDIX

REMOVING ATS-III DIGITIZING NOISE

The ATS-III data are received at the NASA ground station at Rosman, North Carolina. A system block diagram for the spin-scan cloud camera is shown in Fig. A-1 (NASA, 1968). It is evident that the data is received, demultiplexed, sent through a video processor and finally digitized before being recorded on magnetic tape.

Figure A-2 shows more clearly the latter stages of the data flow system. Data from the video processor is sent to three individual A/D converters, one converter for each of the three-color channels—green, red and blue. The data is then simultaneously loaded into a partition core memory and sequentially unloaded onto digital tape.

The A/D converters used in this equipment are Westinghouse High Speed Converter #339D222 (Westinghouse, 1968). This is a successive approximation type converter which has a sequence conversion rate of 10^6 words per second. Westinghouse claims an absolute accuracy of plus 2 bits for this converter. A flow diagram of the operation of this successive approximation converter is shown in Fig. A-3. In this unit a number of decisions are required. The first decision is whether the brightness ample is greater than 128. Depending upon the answer to that decision, the converter next determines whether the brightness sample is greater or less than 64 (or whether it is greater or less than $64 + 128$). This decision process continues sequentially downward through all 8 bits. We have diagrammed only seven of the decision processes, inasmuch as we have used only the higher 7 bits in our examination of the data.

In examining the ATS-III data for January 20, 1968, for a large section of the earth, it was noticed that there were large changes in the frequency of occurrence of one particular brightness value relative to the next value. For example, the frequency of occurrence of brightness value 36 was about one-half as frequent as the brightness value 38. This lack of smoothness in the frequency of occurrence of brightness samples is shown in Fig. A-4. The data for Fig. A-4 were selected from area 38 (of Fig. 18) because that area is predominately cloud covered and should yield a reasonably smooth frequency of occurrence vs. brightness values over a large brightness range. It is also apparent in Fig. A-4 that the same type of problem is present and both the blue

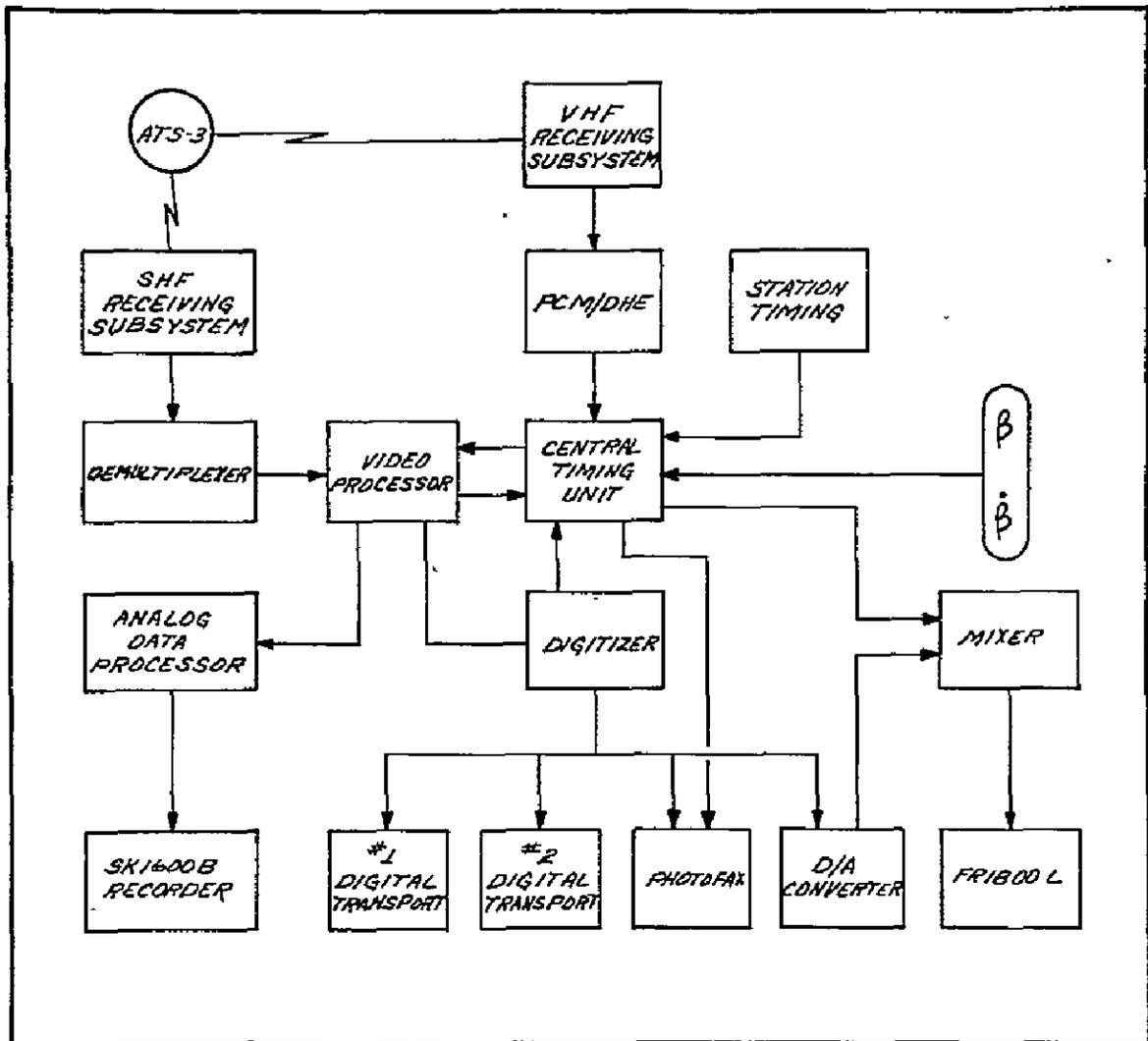


Fig. A-1. Functional Block Diagram of ATS-III Ground Station Data Processing and Recording System.

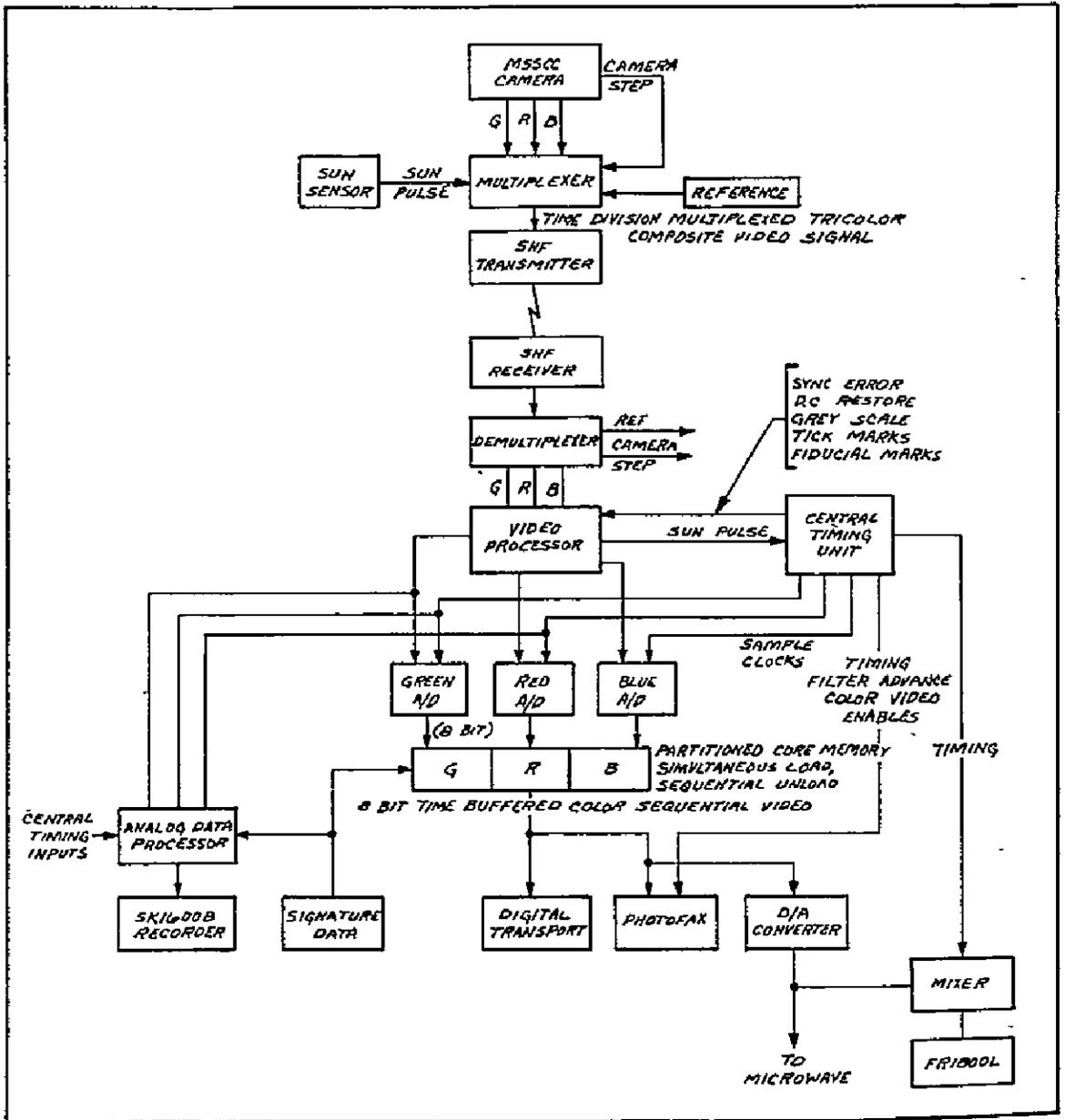


Fig. A-2. Functional Block Diagram of ATS-III Ground Station Data Processing and Recording System.

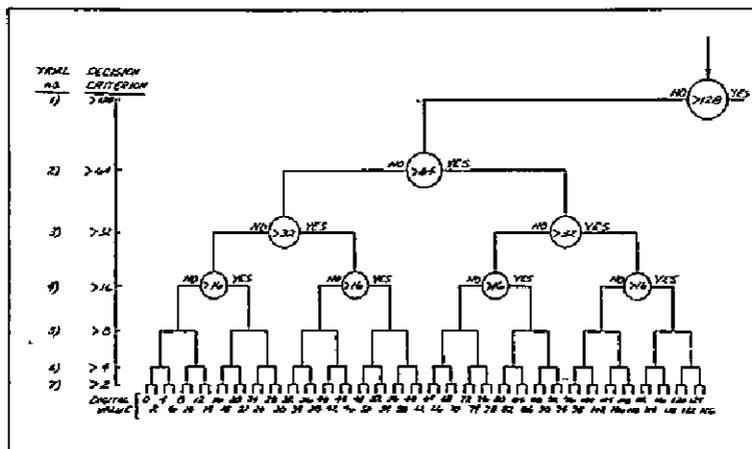


Fig. A-3. Functional Diagram of a Successive Approximation A/D Converter, as used in the ARS-III Data Recording System.

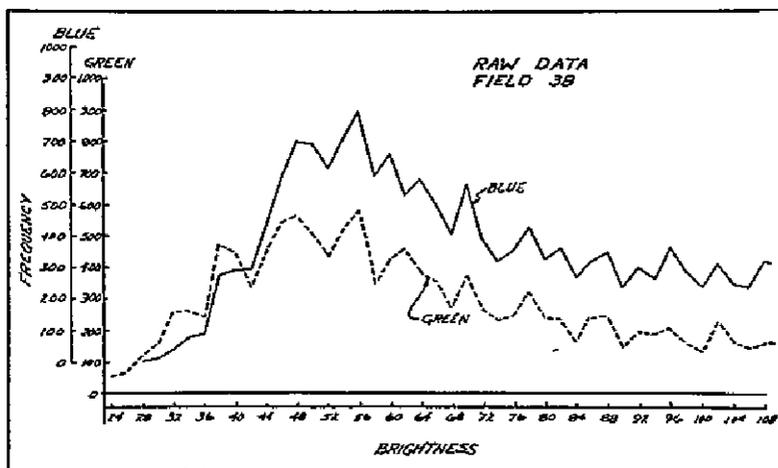


Fig. A-4. Frequency of Occurrence vs. Observed Brightness for Area 38. This shows the data with the Digitizing Error.

and green channels—that is, the same basic pattern of high and low values—occur at the same brightness values in both channels. This strongly suggests it is a characteristic of the digitizing equipment, inasmuch as two identical digitizers were used—one for each channel.

In examining the high and low values in Fig. 18, we have found the error pattern is characteristic of the 2 to 16 bit combinations of the converter. That is, there is a repetitive pattern of error in Fig. A-4 which is related to the 2-16 bit combination. This error pattern is illustrated in Fig. A-5.

In order to correct for this digitizing problem, we have calculated the correction factor required for each individual bit combination shown in Fig. A-5. Smoothed frequency vs. brightness data for areas 38 through 40 (Fig. 18) were used to calculate these correction factors. The corrections are listed in Table A-1. When these corrections are applied, the frequency vs. brightness curve is considerably smoother. For example, field 38 after correction is shown in Fig. A-6. In addition, corrected data for areas 44 and 71 are shown in Figs. 20 and 21.

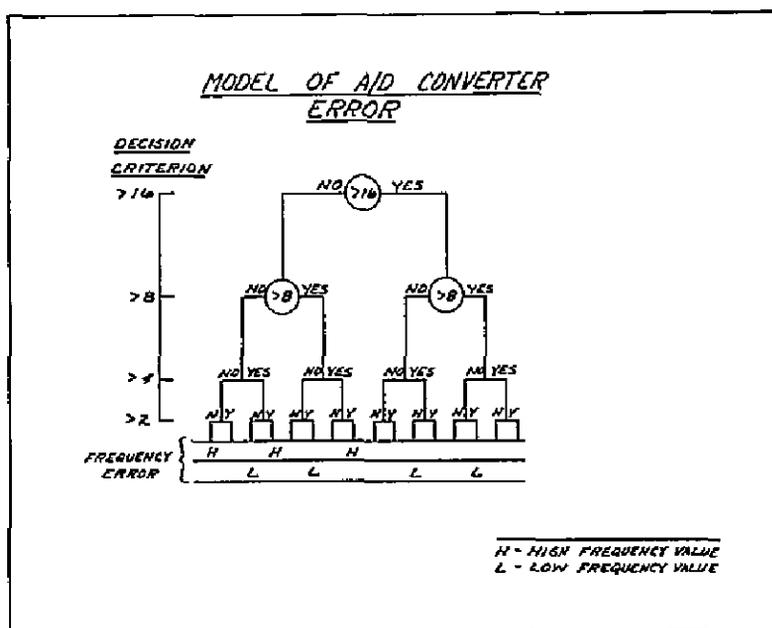


Fig. A-5. Model of A/D Converter Error.

Table A-1

Corrections for Digitizing Error

B	(Brightness Values)			(Channel)	
	B	B	B	Blue	Green
0	32	64	96	0.886	0.900
2	34	66	98	0.992	1.006
4	36	68	100	1.190	1.276
6	38	70	102	0.847	0.797
8	40	72	104	1.071	1.049
10	42	74	106	1.186	1.228
12	44	76	108	1.004	1.003
14	46	78	110	0.884	0.924
16	48	80	112	0.983	0.937
18	50	82	114	1.013	1.011
20	52	84	116	1.163	1.235
22	54	86	118	0.966	0.918
24	56	88	120	0.846	0.823
26	58	90	122	1.173	1.225
28	60	92	124	0.938	0.991
30	62	94	126	1.034	0.933

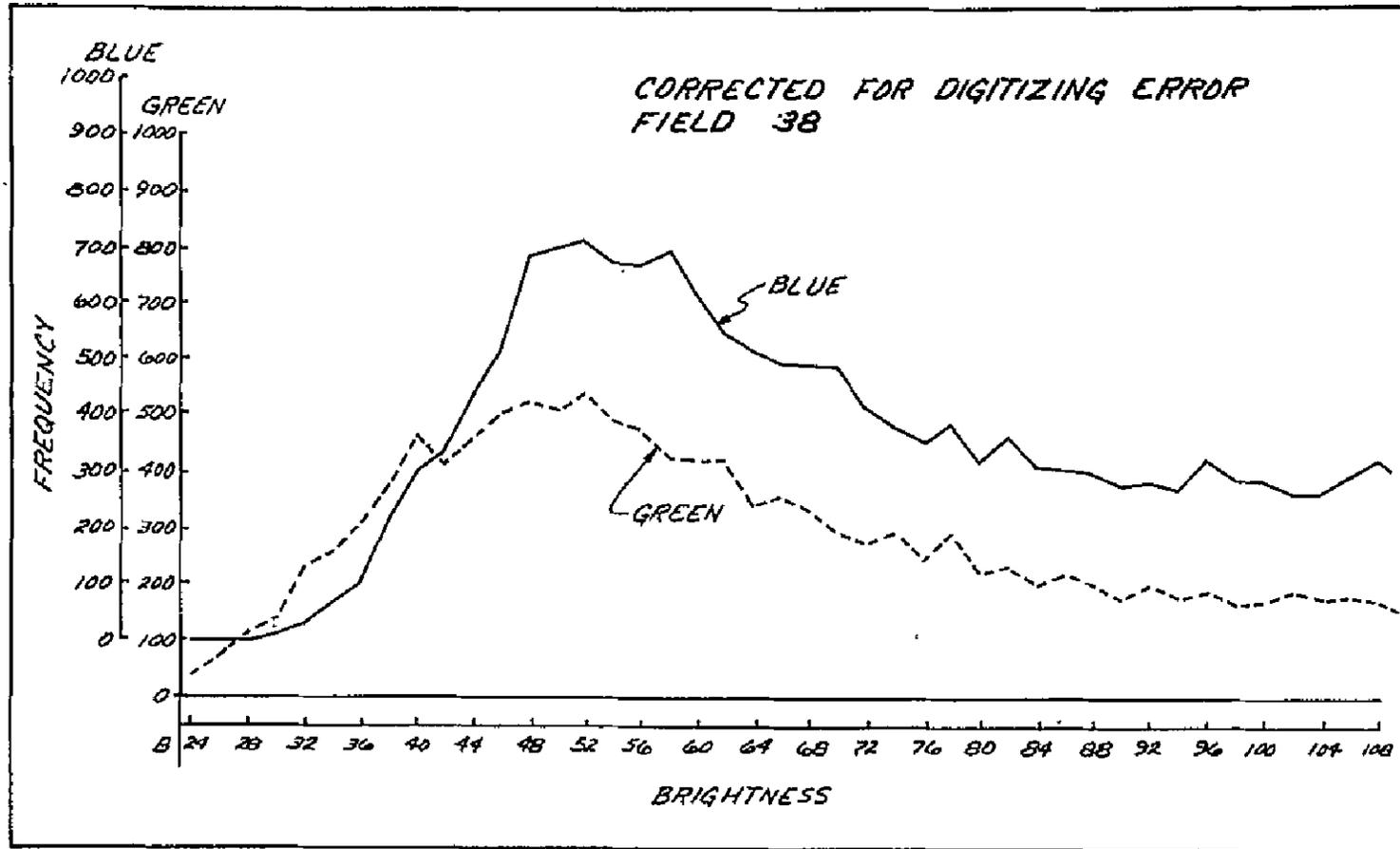


Fig. A-6. Frequency of Occurrence vs. Observed Brightness for Area 38 after the Data have been Corrected for the Digitizing Error.

REFERENCES

- Bartman, F. L., 1967, "The Reflectance and Scattering of Solar Radiation by the Earth." Ph.D. Dissertation. University of Michigan.
- Chandrasekhar, S., and D. Elbert, 1954. "The Illumination and Polarization of the Sunlit Sky on Rayleigh Scattering," Trans. Am. Phil. Soc., New Series, vol. 44, part 6, p. 643.
- Chandrasekhar, S., 1950. Radiative Transfer, Clarendon Press, Oxford, England, 393 pp.
- Coulson, K., 1968. "Effect of Surface Reflection on the Angular and Spectral Distributions of Skylight," Journal of Atmospheric Sciences, 25:5, 759-770.
- Coulson, K., 1967. "Determination of Heights of Cloud Tops by Measurements from Satellites," (unpublished document).
- Coulson, K., J. Dave and Z. Sekera, 1960. Tables Related to Radiation Emerging from a Planetary Atmosphere with Rayleigh Scattering, University of California Press, Berkeley and Los Angeles.
- Dave, J. and R. Warten, 1968. "Program for Computing the Stokes Parameters of the Radiation Emerging from a Plane-Parallel, Nonabsorbing, Rayleigh Atmosphere," Report No. 320-3248, IBM Scientific Center, Palo Alto, California.
- Deirmendjian, D. and Z. Sekera, 1954. "Global Radiation Resulting from Multiple Scattering in a Rayleigh Atmosphere," Tellus, 6:4, 382-398.
- Elterman, L. and R. Toolin, 1965. Atmospheric Optics (Chapter 7), in Handbook of Geophysics and Space Environments, AFCRL-USAF.
- Krinov, E. L., 1947. "Vectorial Reflectance Properties of Natural Formations (Translation)," Laboratoriia Aerometadov, Akad. Nauk. SSSR.
- NASA, 1968, "ATS Project Reports," ATS Project Office, Goddard Space Flight Center, Greenbelt, Md.
- Parent, R., 1969. Private communication.
- Lord Rayleigh, 1899. Scientific Papers (Cambridge, England, 1899), 1, pp. 87, 104, 518.

- Saiedy, F., D. Hilleary and W. Morgan, 1965. "Cloud Top Altitude Measurements from Satellites," Applied Optics, 4: 4, 495.
- Sekera, Z., 1969. Private Communication. The radiance inversion is being completed by a student of Professor Sekera. The polarization inversion has been completed by Sekera, 1967.
- Sekera, Z., 1967. "Determination of Atmospheric Parameters from Measurement of Polarization of Upward Radiation by Satellite or Space Probe," ICARUS 6, 348-359.
- Sekera, Z., and W. Viezee, 1961. "Distribution of the Intensity and Polarization of the Diffusely Reflected Light over a Planetary Disc," The Rand Corporation, R-389-PR; Santa Monica, Calif., 45 p.
- Sekera, Z., 1956. "Recent Developments in the Study of the Polarization of Skylight," Advances in Geophysics, III, pp. 43-104.
- Sekera, Z., and E. V. Ashburn, 1953. "Tables Relating to Rayleigh Scattering of Light in the Atmosphere," NAVORD Sept. 2061, U. S. Naval Ordnance Test Station, Inyokern, Calif., 113 pp.
- Sekera, Z., and G. Blanch, 1952. "Tables Relating to Rayleigh Scattering of Light in the Atmosphere," Sci. Rept. No. 3, Contract AF19(122)-239, Air Force Cambridge Research Center, 85 pp.
- Shettle, E., and J. Weinman, 1969. Private communication.
- Smith, W., 1969. Private communication.
- Superintendent of Documents, Washington, D. C., March 1, 1969. "World Weather Program—Plan for Fiscal Year 1970."
- U. S. Committee for GARP, 1969. "Plan for U. S. Participation in the Global Atmosphere Research Program," National Academy of Sciences, Washington, D. C.
- Wark, D. Q. and D. T. Hilleary, 1969. "A Successful Test of Remote Probing of the Atmospheric Temperature (to be published in Science).
- Wark, D. and D. M. Mercu, 1965. "Absorption in the Atmosphere by the Oxygen A Band," Applied Optics, 4, 839.
- Westinghouse, 1968. "Westinghouse High Speed A/D Converter," Advanced Technology Bulletin No. 6, Westinghouse Defense and Space Center, Baltimore, Md., 21203.

ACKNOWLEDGMENTS

A number of people contributed directly and indirectly to this work.

Professor Verner Suomi has provided stimulation and helpful ideas from the beginning of this research. Professor James Weinman has added much to the final manuscript by his careful and detailed review. Dr. Steve Cox and Dr. Tom Vonder Haar provided many useful discussions during the formative stages of the research. Professor Z. Sekera was very generous of his time in discussions and in reading the thesis. This entire group has been most helpful in contributing their time and ideas, for which I am very grateful.

Professor Kinsell Coulson's (1967) study provided the foundation for the present work and was used extensively in Chapter II. His original ideas concerning the inversion and formulation of the problem are significant parts of the present work and are gratefully acknowledged. Dr. J. Dave most generously contributed his computer program for the calculation of theoretical radiance values. The present work would not have been possible without that useful program.

Mr. Eric Smith did most of the computer programming required for this study. I extend my sincere appreciation for his careful work.

Many others, too numerous to mention, have assisted in a variety of ways. I extend my deep appreciation to each of them for their assistance.

The research was supported by NASA under research contract NAS5-11542.

N71-11616

CONVECTIVE HEAT TRANSPORT OVER THE TROPICAL MID-PACIFIC
AS ESTIMATED FROM A GEOSYNCHRONOUS SATELLITE ALTITUDE

Dhirendra Nath Sikdar

ABSTRACT

This dissertation, to my knowledge, is the first treatment of satellite cloud photographs as the data source for the evaluation of the convective transport of mass and energy from the lower troposphere to the tropical upper troposphere.

In the first part, an objective technique has been developed to estimate the mass and energy exchange in a convection system corresponding to congestus and cumulonimbus intensities. The technique involves measuring the area change of the cirrus outflow on a sequence of satellite cloud photographs obtained at geostationary altitude. The data show (1) that the technique is able to isolate vigorous and moderate convection regimes on the ATS-I and ATS-III satellite photographs, (2) that our values of mass and energy flux are consistent with ground based measurements such as those of Braham (1952) and Brown (1967). Thus, the geostationary satellite photographs can be used to estimate convective mass and heat transport.

The second part deals with the application of the technique to the study of large scale convective mass and energy exchange over the tropical mid-Pacific (15°N to 15°S, 120°W to 180°W). This investigation covers the period April 1 - April 30, 1967 when frequent satellite cloud observations were taken. During this period additional conventional surface and upper air observations were available from the Line Island Experiment.

The results show that the releases of convection on a meso- to sub-synoptic scale over the tropical mid-Pacific are controlled by the large scale motion field and that the convective heat transport pulsates with an approximate periodicity of five days. In the short range time scale (minutes to one day) on some days intense variations in the daytime convective activities are observed, but these variations seem to have no relation with the diurnal heating cycle. Convective activity occurrence is random in character in the short range time scale but is well organized on a time scale of a few days to a month.

Contents

Part 1

	Page
1. A Satellite Estimate of Mesoscale Convective Transport of Mass and Energy	79
1.1 Introduction	79
1.2 A Convection Model for the Tropics	81
1.2.1 Effect of Wind Shear on the Layer of Cirrus Outflow	83
1.3 Mathematical Formulation of the Convection Model	85
1.3.1 Volume Flux	85
1.3.2 Mass Flux	87
1.3.3 Computation of Inflow and Outflow Rates	88
1.3.4 Computation of Latent Heat Flux	89
1.3.5 Compensating Motion in the Cloud Environment	89
1.4 Data Source and Technique of Analysis	90
1.5 Interpretation of ATS-I Satellite Cloud Photographs	96
1.6 Comparison with Other Models	105
1.6.1 Thermodynamic State over Palmyra on April 19, 1967	106
1.6.2 Mass Flux Data	109
1.6.3 Energy Flux Data and the Cloud Mass	112

Part 2

2. A Satellite Estimate of Large Scale Convective Transport of Mass and Energy	114
2.1 Introduction	114
2.2 Implications on the "Hot Tower" Concept	115
2.3 Grid Selected in this Study	116
2.4 ATS-I Satellite Data Coverage	117
2.5 Technique of Analysis	117
2.6 Vertical Cross-sections of Temperature and Moisture in the Grid Box	122
2.7 Short Range Fluctuation of Meteorological Events in the N. H. and S. H. Sectors	127
2.7.1 Percent-bright Cloud Cover	127
2.7.2 Average Upward Energy Transport	127
2.8 Long Range Fluctuations of Meteorological Events in the N. H. and S. H. Sectors for April 1967	127
2.8.1 Cloudiness Cycle	127
2.8.2 Daily Mean of the Average Positive Cloud Divergences	132

2.8.3	Mass Inflow Rate at the Subcloud Layer and the Convective Heat Transport to the Tropical Upper Troposphere during April 1967	132
2.8.4	Phase and Speed of the Upper Troposphere Disturbance Field	136
2.8.5	Implications for Modeling the Tropical Atmosphere	136
3.	Summary	137
4.	Proposed Further Investigation with the ATS-I Photographs	138
5.	References	139
	Appendix I	143
	Appendix II	151

1. A SATELLITE ESTIMATE OF MESOSCALE CONVECTIVE TRANSPORT OF MASS AND ENERGY

1.1 Introduction

The first section of this paper is an attempt to show that, subject to some reasonable assumptions, one can estimate the vertical heat transfer over the tropics from the ATS-I cloud photographs.

It has long been recognized that the tropics are the heat sources, while the polar regions are the heat sink, and that over a long time scale a balance exists. Most of the heat is transferred upward and poleward from the tropics through convective systems of various scales. A number of studies on the dynamics of a single convective element have been done: Lilly (1962), Ogura (1963), Orville (1965), Asai and Kasahara (1967). In the numerical prediction model, one need not include the effect of a single convective element on the large scale flow pattern. It is enough to incorporate the dynamical interactions of a convective ensemble on the large scale circulations, taking into the consideration the feed-back mechanisms that are presumed responsible for the organization of cumulus convection in the tropics.

There are three important factors involved in the investigation of clouds. First, the released latent heat in clouds is a major source of heat for the atmosphere. Second, large amounts of sensible heat, water vapor and momentum are transported by cumulus convection. Third, the presence of clouds modifies

and redistributes the heat sources and sinks in the earth-atmosphere systems. In numerical prediction models, at present, the effects of latent heat release are incorporated by predicting a hydrological cycle. The second effect is simulated by various processes called "convective" adjustments. Here again, the modeler looks for the composite effect of two causes in the convective systems so as to simplify the prediction model. The third effect is taken into account by a climatological distribution of clouds. So, if one desired to include the prediction of cloud population and its interaction to the large scale motion field in the numerical model for general circulation, a knowledge of the spatial distribution and magnitude of heat sources and sinks with time is an important requirement.

So far the effects of the mesoscale process on the large scale circulation are not incorporated in the numerical prediction model because their interactions are not well understood yet. Only in recent years, efforts have been made to parameterize the small scale motions with respect to the general circulation. Byers and Braham (1949) have noted remarkable dynamical characteristics of the mesoscale convective systems from the thunderstorm project. Malkus (1955) has studied the organization of the tropical cumuli related to large scale circulations. Many interesting structures of mesoscale systems have been revealed in the past decade (Fujita 1963, Newton 1963).

To enable one to study the mesoscale systems over a wide area such as the tropics, one needs a higher density of observation network having spatial separation between two stations not exceeding 2 km. This is not feasible, especially over the ocean. Therefore, for such investigations, one has to resort to aircraft reconnaissance observations (Malkus and Riehl, 1964) which are very expensive and therefore scarce. A meteorological satellite, on the other hand, has immense potential when it is equipped with sensors capable of sensing the cloud systems of dimensions less than 2 km and has the sampling frequency less than the life cycle of a convective ensemble. Satellite studies on cloud processes commenced as early as 1960 with the TIROS series of satellite sensors, but the crux of the solution came in sight with the launching of geosynchronous satellite carrying a spin-scan camera over the Pacific equator in December 1966. For the first time, meteorologists could view the growth and decay of the cloud systems over the Pacific region in spatial scale ranging from 2.3 nautical miles to the planetary scale. As regards the data on the cloud population and the scales of convection over the tropics, one started receiving volumes of information on these geosynchronous satellite photographs. The problem remains as to how best this information can be utilized for better understanding of the tropics which is the fuel source for the atmospheric heat engine.

1.2 Convection Models in the Tropics

In the tropics, convection systems may be classified into two broad classes: (a) shallow convection, and (b) penetrative convection. Over the tropical oceans, shallow convection is associated with cumulus clouds whose tops do not penetrate the tropical inversion layer, while penetrative convection extends to the upper troposphere to at least 500 mb. As pointed out by Riehl and Malkus (1958), this penetrative convection may be the prime mechanism in the vertical transport of heat.

Columnar convection is the penetrative type of vertical motion wherein less dense fluid streams upward in continuous columns into more dense fluid. Byers and Braham (1949) found that measurements in the large scale cumulus clouds and thunderstorms indicated updraft throughout most of the cloud heights. The concept of thunderstorm as a mesoscale phenomenon is very easily seen on the radar scope where individual echo cells do not distribute uniformly over a wide range of synoptic scale disturbance, but often concentrate into an organized pattern. Vonnegut, Moore and Botka (1950), in their experimental findings, supported the concept of deep connected flow in thunderstorms. It may be emphasized at this stage that the divergence associated with the mesoscale system is of the order of 10^{-4} sec^{-1} . This is one order larger than that of the synoptic field. According to Malkus et al. (1964), the synoptic field presumably provides a favorable condition for the development of convections.

Ackerman (1959), in her study on the liquid water distribution in the tropical cumuli, has established a positive correlation between the liquid water content and the cloud width and depths. The increases in observed and relative water contents would imply that the net effect of the entrainment of clear air is less in large clouds than in small ones. In large clouds, the lateral entrainment at the cloud periphery may be less effective in diluting the core because of the time necessary for diffusion into the cloud interior by mixing. Her observation supports further that the undiluted or nearly so penetrative convection might be a common feature in a tropical convective cloud ensemble that is responsible for the vertical transport of mass, momentum and heat.

Figure 1 is a graph of entrainment rate $\frac{1}{M} \frac{dM}{dz}$ as defined by Stommel (1947) as a function of convective cloud diameter D in Km for different values of exchange coefficients K (Malkus, 1960). The variations of K seem to make a minor difference in dilution compared to the effect of cloud size. According to Levine (1959), the value of $K = 0.5$ is emphasized to be more reasonable for large cumulus towers. This diagram indicates that under saturation conditions, which are obvious in the tropical lower troposphere over the Pacific, entrainment rates of less than $0.1 \times 10^{-5} \text{ cm}^{-1}$ lead to insignificant dilution of clouds. In our analysis most of the clouds observed by satellite are much larger than 10 km in diameter. Therefore a few of the large towers in a large convective cloud complex will have no difficulty in reaching the

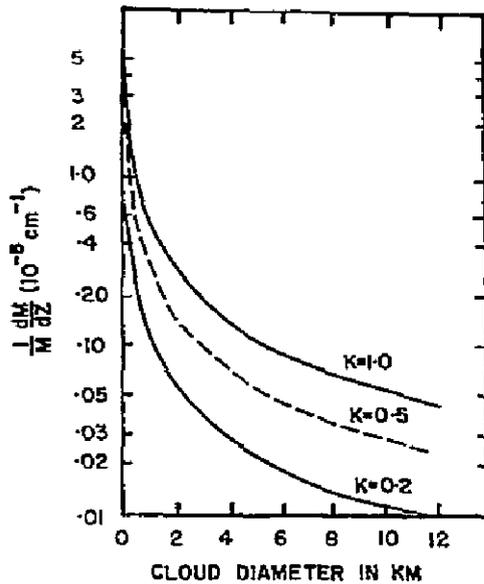


Fig. 1. Entrainment rate $\frac{1}{M} \frac{dM}{dz}$ as a function of convective cloud diameters.

tropopause and spreading laterally from there.

Malkus, Ronne and Chaffe (1961), in a photogramatic study of the distribution, found that convergence of towering cumuli in the central part of a hurricane is several percent. This would mean that the development of the cumulus clouds is a dynamic necessity for the most efficient heat transport upward in a hurricane.

The convection model we propose in this paper is an open flow system consisting of three sections of flow: (a) inflow at the sub-cloud layer; (b) vertical motion in the cloud cross-section; and (c) outflow near the tropopause. The assumption of a steady state condition simplifies our flow model. A schematic diagram showing the flow systems in the model is presented in Fig. 2.

(a) The inflow layer, as shown in Fig. 2, is the layer between the sea surface and the lifting condensation level. In this layer, warm moisture laden air is drawn into the cloud region because of the locally produced thermal convergence in the surface layer. This layer also contains a part of the downdraft descending below the LCL in the mature and dissipating stage of convection. The thermal stratification and the moisture distribution in the warm air inflow would virtually decide on the degree of evaporation in the cold downdraft that enhances the sensible heat transfer from the subcloud layer. In the model, the lowest layer in all probability is considered to be the heat source throughout the life cycle of convection.

(b) The second layer is the convection layer extending from the LCL to the "throat" (Bates, 1961), of the cloud near the cloud top where the buoyancy force is zero or slightly negative. In this layer the vertical motion will be maximum in the core, decreasing to a negligible value near the cloud periphery. In the case of the symmetric updraft regime, one could imagine this decrease of vertical speed to be linear across a horizontal cross-section of the cloud.

Let us assume that the cumulus clouds under consideration are cylindrical in shape as shown in Fig. 2 and consist of an inner core undiluted by entrainment. These clouds are axisymmetric such that it acts like a cylindrical pipe pumping the fluid from the subcloud medium. It is conceivable that the warm air inflow in the subcloud layer rises through the cylinder after converging about a point.

One assumes that the buoyant air starts condensing above the LCL where the latent heat of condensation is released. An adiabatic model in general presumes that a unit volume of air while rising upward expands at a rate depending on the pressure and temperature distribution in the vertical, and the latent heat, released through condensation or sublimation, is conserved in the volume itself. This feature is also schematically shown in Fig. 2. The release of latent heat increases the buoyancy force of the rising parcel, and as a result, more water vapor condenses out. From the Thunderstorm Project (Byers and Braham, 1949), we find that the stronger the updraft speed, the larger is the amount of liquid water carried through a greater depth. This might be the reason for a substantial amount of liquid water to sublimate into ice crystals in towering cumuli. The outflow of these ice crystals through the top of the tower appear as a shield in the form of an "anvil." The adiabatic model gives rise to an increasing vertical velocity with depth to a level and then decreases above that level.

(c) The outflow layer extends from the throat of the convective tower to the base of the tropopause. It is assumed that faster rising cloud particles lose their vertical momentum when they arrive very close to the throat of the tower and therefore they are under the influence of local wind. This damping effect of the stable layer below the tropopause on the vertical motion field has long been recognized. The situation is similar to a jet of air particles impinging on a flat plate. The air reaching to the top of the storm comprises the source for divergence in the cloud plume at the top.

1.2.1 Effect of Wind Shear on the Layer of Cirrus Outflow

Figure 2a shows the model outflow in the absence of wind shear in the upper troposphere while Fig. 2b presents the outflow in the presence of pronounced wind shear in the upper troposphere. The model circulations in these two cases are also presented.

In the former case, for the presence of cloud material all through the cloud depth, the reflectance at the cloud top is expected to be very strong. In the latter case the cirrus shield will be blown off from the cumulonimbus tower and be embedded in the dry atmosphere. The reflectance will be significantly less than in the former case for the part of the incident solar energy will transmit through the thin cloud material and get absorbed in the lower atmosphere. In the present study from the cloud reflectance analysis over the large scale cloud field, we had no occasion to identify such detached cirrus shields. The wind shear may change the shape of the cloud top but as long as the cloud thickness remains the same our flux computation scheme will be valid.

The outlines of this simple convection model indicate the possibility of probing into mesoscale characteristics as viewed from a stationary platform in space. The geosynchronous satellite views the lateral extent of the cloud field at different time intervals. If one is able to identify the types of convective clouds from a satellite photograph, it would be possible to evaluate the

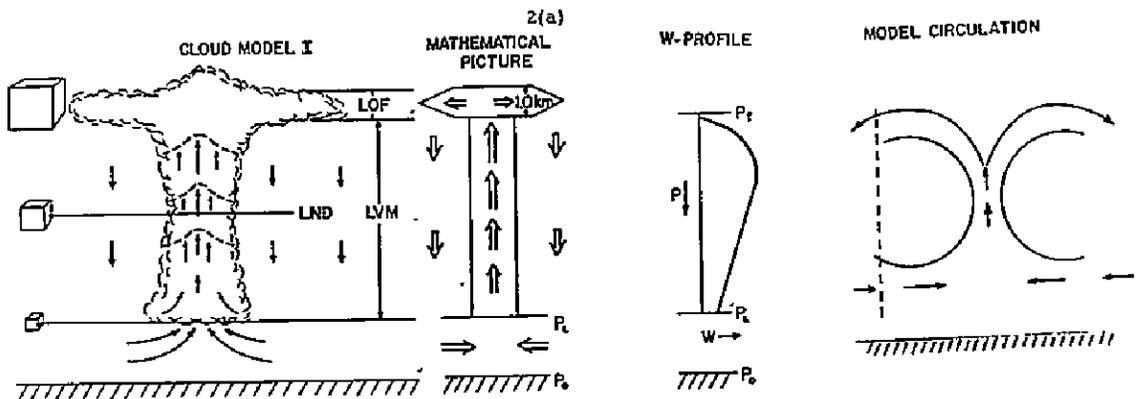
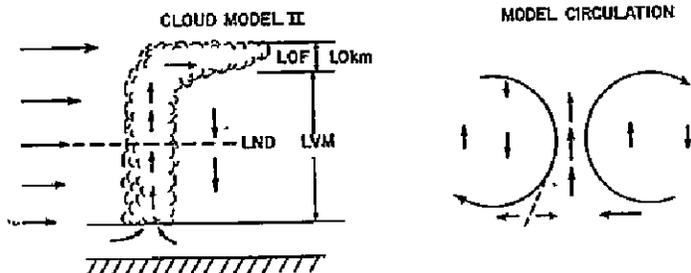


Fig 2b WIND SHEAR AT THE UPPER TROPOSPHERE:



SCHEMATIC DIAGRAM FOR THE PROPOSED CONVECTION MODEL

Figs. 2a, b. Schematic diagram of a tropical convection model.

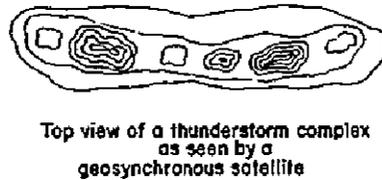
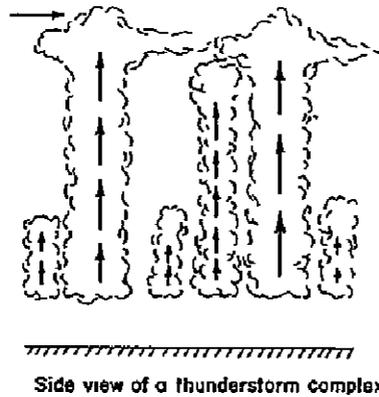


Fig. 2c. Schematic diagram of a convective complex as viewed by a geosynchronous satellite.

rate of growth and decay of a convective ensemble from the successive time lapse photographs. The technique that has been used to classify such convection systems is presented in the latter part of this paper.

1.3 Mathematical Formulation of the Model

In Sec. 1.2, we have shown that the convective cloud with diameter larger than 10 km has negligible entrainment. Thus, in the following mathematical formulation we will assume that the mass is conserved in the second layer (LVM) of the model.

1.3.1 Volume Flux

The volume at the level of divergence near the cloud top may be related to the volume at the subcloud layer, using the equation of state, by

$$V_t = \left(\frac{P_L T_t}{P_t T_L} \right) V_L \quad (1)$$

where suffix "t" and "L" stand for the base of the outflow layer and top of the inflow layer, respectively; T_t and T_L are the corresponding virtual temperatures.

In case of ideal flow through the cylinder, P_t represents the static pressure at the top and should be small. In the model P_t is defined as the lowest pressure to which a buoyant cloud will mainly penetrate from the surface layer. In most of the tropical soundings investigated in this work from parcel consideration, P_t has been found to be close to 150 mb. Updraft air reaches an equilibrium at P_t and then spreads out in a layer probably not more than 1 km thick (Ludlum, 1966). This flow model presumes that the cirrus outflow at the top of the cloud tower remains confined to a layer of thickness 1.0 km (Borovikov, 1963).

Taking the mean value for P_t , P_L , T_t , and T_L , the volume flux $\frac{dV_t}{dt}$ at the layer of outflow may be related to the volume flux at the layer of inflow, by the equation

$$\frac{dV_t}{dt} = \left(\frac{P_L T_t}{P_t T_L} \right) \frac{dV_L}{dt}$$

Writing $V_L = \frac{M_L}{\rho_L}$, we get

$$\frac{dV_t}{dt} = \left(\frac{P_L T_t}{P_t T_L} \right) \frac{d}{dt} \left(\frac{M_L}{\rho_L} \right) \quad (2)$$

where ρ_L is the density of air at the subcloud layer. For a steady flow at the subcloud layer, ρ_L may be assumed independent of time in which case equation (2) may be written as

$$\frac{dV_t}{dt} = \left(\frac{P_L T_t}{P_t T_L \rho_L} \right) \frac{dM_L}{dt} \quad (3)$$

In equation (3), P_t , P_L and T_t are evaluated from the mean upper air sounding in the vicinity of the cloud ensemble under investigation. The term dV_t/dt , however, is computed in an indirect method from the ATS-I satellite cloud photographs.

1.3.2 Mass Flux

The amount of mass transported vertically into the layer of outflow (LOF) through the convective towers in time Δt is given by $\frac{\partial}{\partial z}(\rho w)\Delta t$. In this model, we have assumed a steady state flow, and also the cloud motion represents the wind field in the cirrus shield. The term

$\frac{\partial \rho}{\partial t} + \vec{V} \cdot \nabla \rho$ in the equation of continuity may be neglected for the following reasons:

- (1) The local change of density in the assumed steady state flow may not be significant over a period of three photographs or more (i. e., more than an hour).
- (2) In a large convective cloud, the advection of density into the core is assumed negligible.

Under these assumptions one is able to establish

$$\left(\frac{\partial u_c}{\partial x} + \frac{\partial v_c}{\partial y} \right) = - \frac{\partial w_c}{\partial z} \quad (4)$$

where subscript "c" refers to cloud particles. The quantity

$$\left(\frac{\partial u_c}{\partial x} + \frac{\partial v_c}{\partial y} \right) = \nabla_H \cdot \vec{V}_c$$

can be related to increase or decrease of the area encompassed by the curve of cloud particles, and does not necessarily reflect on the actual horizontal flow divergence $\nabla_H \cdot \vec{V}$.

Let us consider a cloud volume $V = dx \cdot dy \cdot dz$ at time t_0 . After a small finite time Δt , the cloud particles bounding the area will have moved and the new volume would be

$$\begin{aligned} & \left(dx + \frac{\partial u_c}{\partial x} dx \Delta t \right) \left(dy + \frac{\partial v_c}{\partial y} dy \Delta t \right) dz \\ & = dx dy dz \left(1 + \left(\frac{\partial u_c}{\partial x} + \frac{\partial v_c}{\partial y} \right) \Delta t + \left(\frac{\partial u_c}{\partial x} + \frac{\partial v_c}{\partial y} \right) (\Delta t)^2 \right) \end{aligned} \quad (5)$$

since dz is assumed constant.

¹This cloud divergence can be considered analogous to the horizontal spread of smoke particles in the atmosphere from a smoke chimney which occurs even in the absence of flow divergence, such as in a constant flow field.

The change in cloud volume is

$$\begin{aligned} \Delta V &= dx dy dz \left(\left(\frac{\partial u_c}{\partial x} + \frac{\partial v_c}{\partial y} \right) t + \frac{\partial u_c}{\partial c} \frac{\partial v_c}{\partial y} (\Delta t^2) \right) \\ &= V \left(\nabla_H \cdot \vec{V}_c + \frac{\partial u_c}{\partial x} \frac{\partial v_c}{\partial y} (\Delta t) \right) \Delta t \end{aligned} \quad (6)$$

since $V = dx dy dz$

$$\nabla_H \cdot \vec{V}_c = \frac{1}{V} \frac{dV}{dt} \quad (7)$$

in the limit Δt approaching zero.

This would mean a fractional change in volume flux of cloud particles around a point is equal to the radial divergence. With dz as constant, the equation (7) may be written as

$$\nabla_H \cdot \vec{V}_c = \frac{1}{A_c} \frac{dA_c}{dt} \quad (8)$$

where $A_c = dx dy = \pi r_e^2$, r_e is the equivalent radius of the cirrus shield. From equation (8), we get

$$\nabla_H \cdot \vec{V}_c = \frac{2}{r_e} \frac{dr_e}{dt} \quad (9)$$

1.3.3 Computation of Inflow and Outflow Rate

The anvil may be regarded as resulting from a circular disk expanding uniformly normal to its edge. The outflow rate in the layer of 1.0 km thickness (Ludlum, 1966; Borovikov, 1963; Anderson, 1960) is given by

$$\frac{dV_t}{dt} = ((\pi r_e^2) \cdot dz \cdot (\nabla_H \cdot \vec{V}_c)) \text{ cm}^3/\text{sec} \quad (10)$$

where $dz = 1.0$ km.

Combining equations (3) and (10), one obtains the mass inflow rate at the base of the cloud tower as

$$\frac{dM_L}{dt} = \frac{P_t T_L P_L}{P_L T_t} ((\pi r_e^2) \cdot dz \cdot (\nabla_H \cdot \vec{V}_c)) \quad (11)$$

where $\rho_L = 1.1 \times 10^{-3} \text{ g/cm}^3$.

The term $((\pi r_e^2)(\nabla_H \cdot \vec{V}_C))$ in equation (13) is computed from the measurements on the successive ATS-I satellite cloud photographs.

1.3.4 Computation of Latent Heat Flux

In this model it is assumed that the equivalent potential temperature at the LCL is constant for a saturated ascent and that all water vapor passes into the condensed phase during its ascent through the cloud core. From the upper air sounding data, one can estimate the amount of moisture likely to condense out in saturated adiabatic ascent, let this be $W = (W_L - W_t) \text{ g/kg}$. The latent heat flux is

$$E_c = L \cdot \Delta W \cdot \left(\frac{dM_L}{dt}\right) \text{ cal/sec} \quad (12)$$

where $L = 597 \text{ cal/gm}$ at 0°C .

1.3.5 Compensating Motion in the Cloud Environment

The law of mass continuity requires that the mass lifted up to the upper troposphere through the convective cores should be compensated by the downward motions somewhere in the atmosphere. One can conceive of such motions to exist in the near and far regions of the penetrative convection zones. The far region compensating motions merge with the sinking motion associated with the permanent centers of action while the near region compensating motion exists in the vicinity of the cirrus outflow.

The flow in deep convection has been shown schematically in Figs. 2a, b. In these diagrams the arms of the descending circulation have been left open to show that in the large scale flow field the descent continues in other systems and that there is no boundary on that side (see Green, Ludlum and McIlveen, 1966). The re-entry of the far region descending air in the surface layer of the same latitude takes a long period of up to 20 days or more. We will assume that most of the return flow occurs at a distance from the cloud.

It is probable that the near region compensating motion would evaporate some of the ice crystals in the cirrus shield and as such the area of the cirrus shield as viewed by a geosynchronous satellite would be less than expected. But, Darkow (1963) has shown that these ice crystals in the tropical upper troposphere have much longer life time than that of a convective cloud. In view of his findings, it is presumed that there will be no significant error in the mass and heat flux estimates due to evaporation of ice crystals.

1.4 Data Source and the Technique of Analysis

From the ATS-I cloud photographs and digital information, we are able to do cloud sampling in time and space at approximately 20-22 minutes to a year, and 2.3 nautical miles to planetary scales, respectively. The grid matrix of the photographs consists of 8196 digital elements and 2018 scan lines. On this grid, we have information on brightness ranging from 0 to 256 digital numbers. The advantage of the digital display is that the outlines of small scale convections are difficult to draw on the satellite photographs, but they can very well be distinguished on the digital display. The large scale convective systems, however, can be analyzed from the enlarged photograph itself.

The satellite camera observes reflected solar radiation in the wave band 4500 to 6500 Å. A solar spectrum is fairly flat in this region of the spectrum. A cloud target would appear brighter when its albedo is greater and also when the cloud top is high, one will have a small amount of attenuating atmosphere above it.

Also, the brightness of a cloud top depends on the solar zenith angle. A cloud surface is not a Lambertian reflector, especially at large zenith angle of the sun (Bartman, 1967). However, for small zenith angle ($+30^\circ$ to -30°) of the sun, the top surface of a deep cloud may closely behave as a Lambertian reflector. Figure 3 is a plot of a set of brightness values as a function of time at 5°N , declination 10.9° . The brightness value hardly changes by a few percent in the two hour period around local noon.

Our primary objective is to measure the cloud area change in large clouds on the successive ATS-I digital displays. The cloud area depends on the choice of the brightness level but is independent of the brightness gradient. So, one has to be careful in selecting the brightness cut-off level and the period of observation. In order to avoid corrections to the brightness values, we have chosen satellite pictures close to local noon at the subsatellite point. From scanning many digital displays, it is noted that even the brightest clouds do not exhibit brightness values beyond 140-150 digital count, except on occasions when the return signal is from a specular reflection zone or from a region of intense thermal stratification such as hurricanes, cyclones and convective systems embedded in the ITCZ. These features are, however, well reflected on the brightness gradient. A schematic diagram is presented in Fig. 2c showing the nature of the brightness distribution as would be expected of a thunderstorm complex viewed by a geosynchronous satellite.

The digital information needed is cropped from the original digital tape using the B5500 computer and then displayed in a desired grid through the CDC 3600. Since the signal intensity varies from target to target, there is present a wide range of variations in the digital records. The presence or absence of a cloud over the sea surface can be recognized from the variations in the digital count. The calm sea surface will have maximum absorption of solar radiation

in a cloudless atmosphere and its albedo would be low enough to be distinguished from the cloud surface. A plot of digital signal across the cloud complex is helpful at this point (see Fig. 4). By selecting proper digital levels in the cropped records, we can easily isolate the cloud area. Since the luminance of a particular cloud sample depends on the cloud properties like droplet size spectrum, liquid water content, etc., the study of the brightness range and its gradient in the cloud population should reflect on these characteristics, at least qualitatively. Here we assume that clouds of the same type have approximately the same material characteristics at the top surface.

Figure 4 shows the nature of the signal level across a convective ensemble. It is clear from this diagram that the signals higher than 75 digital counts probably correspond to the regions of deeper clouds in this convective ensemble. In this paper we will present two case studies—one relating the penetrative convection in a small scale cloud complex and the other in a large scale complex.

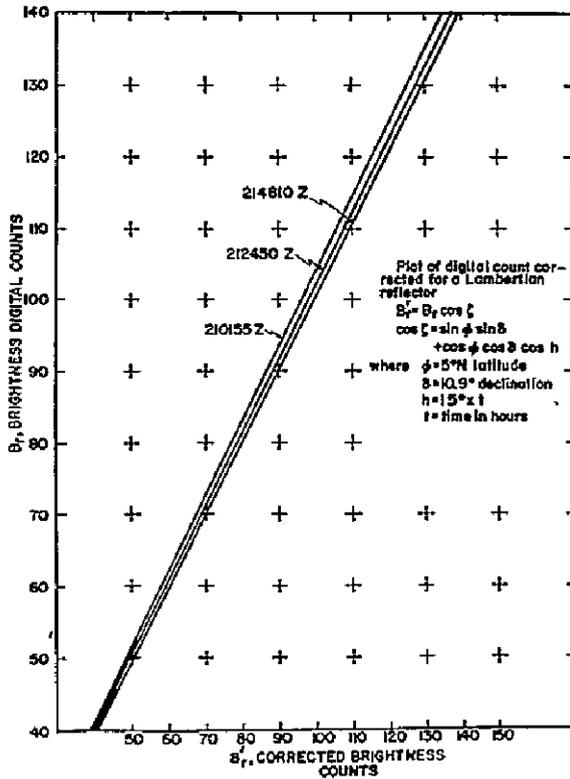


Fig. 3. Plot of corrected brightness values at different solar zenith angles at the sub-satellite point.

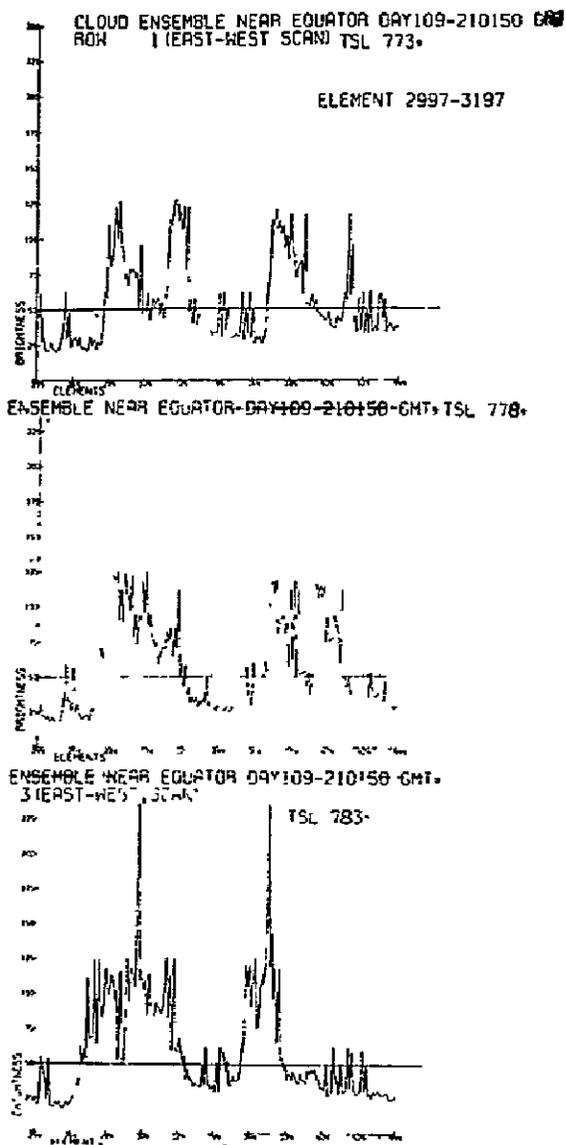


Fig. 4. A plot of brightness values along a scan line across a convective ensemble.

In this study, we have selected the day April 19, 1967, when small scale cumulus activities are seen in the regions of the Line Islands. The mesoscale convective activities on the northwest sector of Palmyra ($5^{\circ}53'N$, $162^{\circ}05'W$) are supported by aircraft reports as well as radar observations close to the period of satellite observations. Figures 5a, b, c represent the ATS-I cloud photographs taken at 210155Z, 212450Z, 214810Z, respectively, on April 19, 1967. Figures 6a, b, c give the enlargements of a section of the bordered areas of Figs. 5a, b, c, respectively. Figures 7a, b, c are the enlargements of a section of cloud field on the ITCZ in the southern hemisphere (location: $10^{\circ}S$, approximately). Figures 8a, b, c and 9a, b, c represent the digital displays for the bordered areas in Figs. 6a, b, c and 7a, b, c, respectively. The object is to show the brightness contrast in the two types of convection regime. Figure 7 corresponds to a moderate convection regime while Fig. 8 is an intense one associated with the southern ITCZ.

The areas of the clouds outlined by the digital number 50 in Figs. 8a, b, c and by the digital count 75 in Figs. 9a, b, c are measured by a planimeter. In the computation of the scale factors on these displays, the distance between the two lines and two elements have been taken at 2.3 nautical miles and 0.7 nautical miles, respectively.

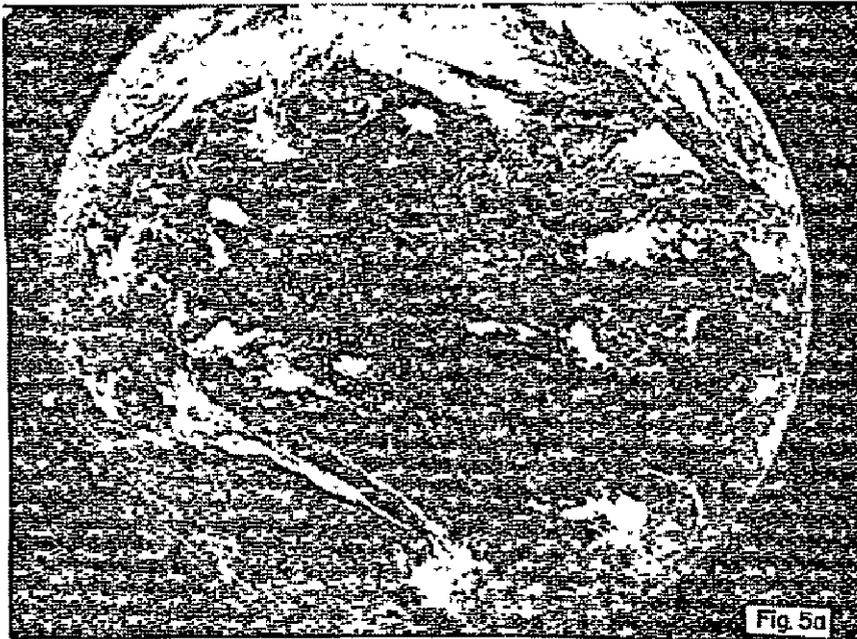


Fig. 5a. ATS-I satellite cloud photograph on April 19, 1967 at 2101 GMT.

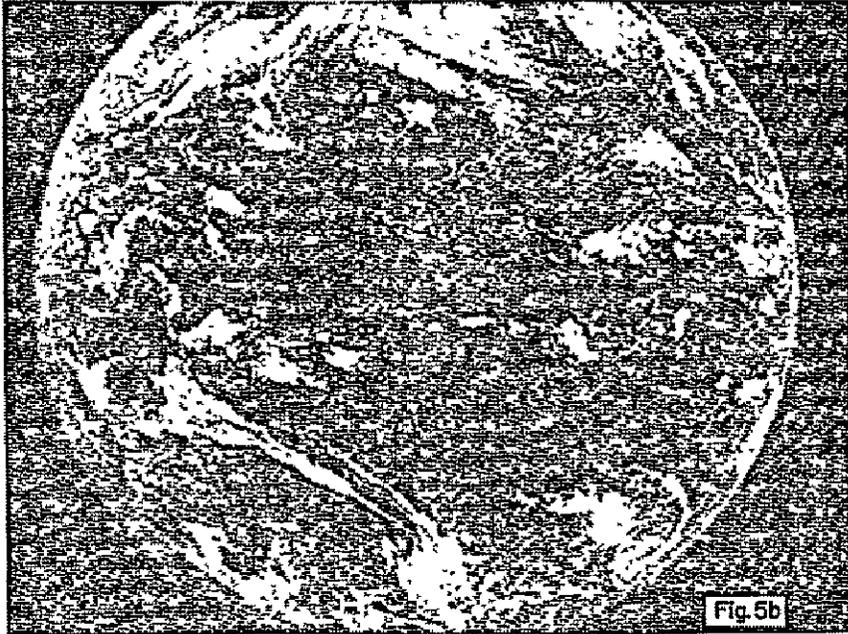


Fig. 5b. ATS-I satellite cloud photograph on April 19, 1967 at 2124 GMT.

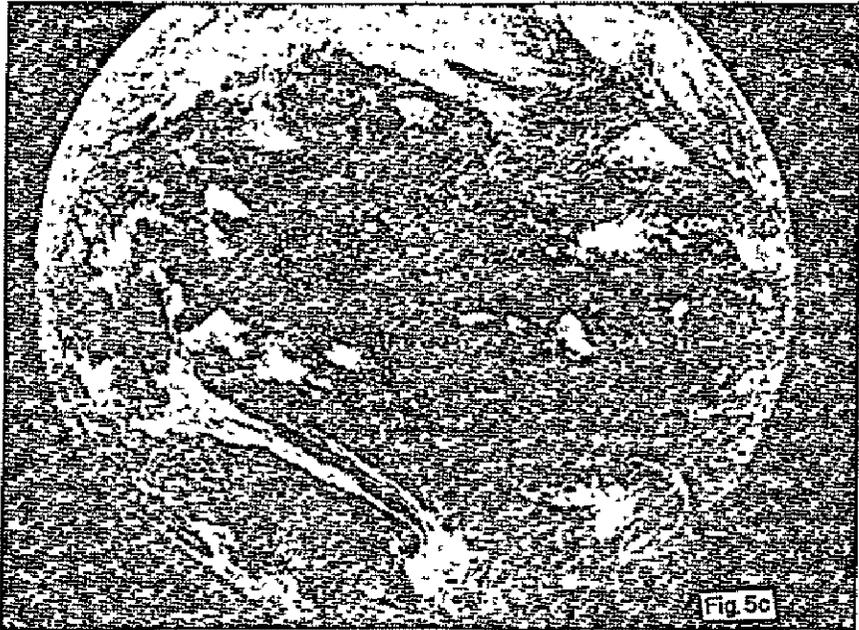


Fig. 5c. ATS-I satellite cloud photograph on April 19, 1967 at 2148 GMT.



Fig. 6a. Enlargement of a section of cloud field around Palmyra at 2101 GMT on April 19, 1967.

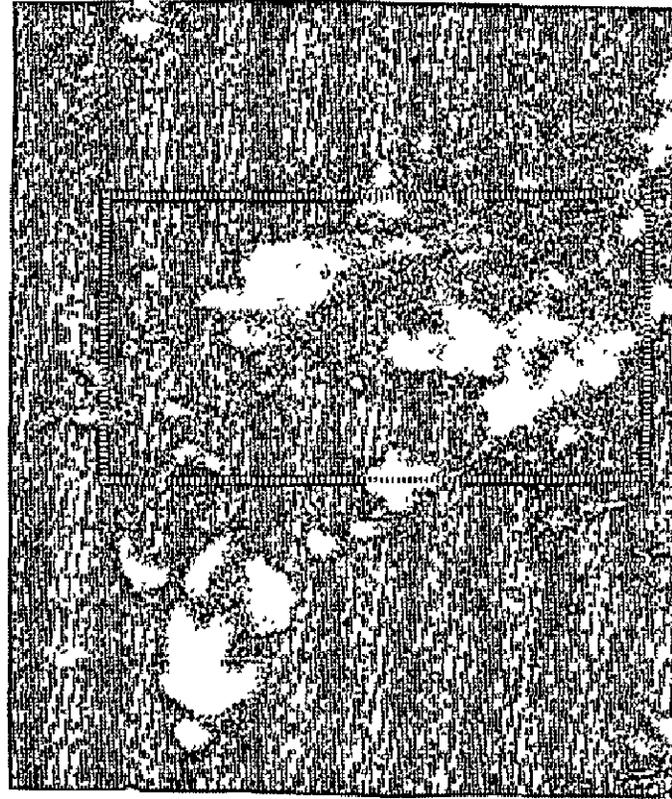


Fig. 6b. Enlargement of a section of cloud field around Palmyra at 2124 GMT on April 19, 1967.

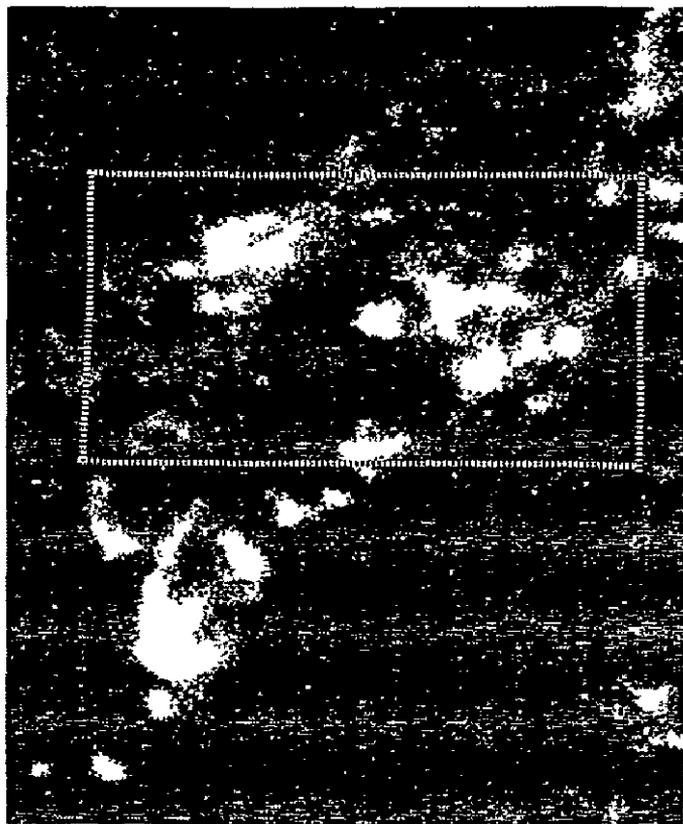


Fig. 6c. Enlargement of a section of cloud field around Palmyra at 2148 GMT on April, 1967.

1.5 Interpretation of ATS-I Satellite Photographs

The first step in the interpretation of ATS-I satellite photographs is to identify the different cloud forms, their brightness values, pattern, structure and size. The brightness factor plays an important role in the cloud interpretation.

Whitney (1963) points out certain identifying characteristics of the satellite observed thunderstorm complexes, i. e., rounded, very bright appearance, large scale, and frequent peripheral clear zones which help to isolate thunderstorm complexes from other cloud complexes.

It is obvious that the greater brightness from a cloud surface is expected to be associated with clouds having greater depth as the solar energy reflected from these cloud tops will have lesser atmospheric attenuation and also depends on the reflectance property of the surface (Conover, 1962). For low clouds, on the other hand, the solar energy (visible) has to traverse a greater path length

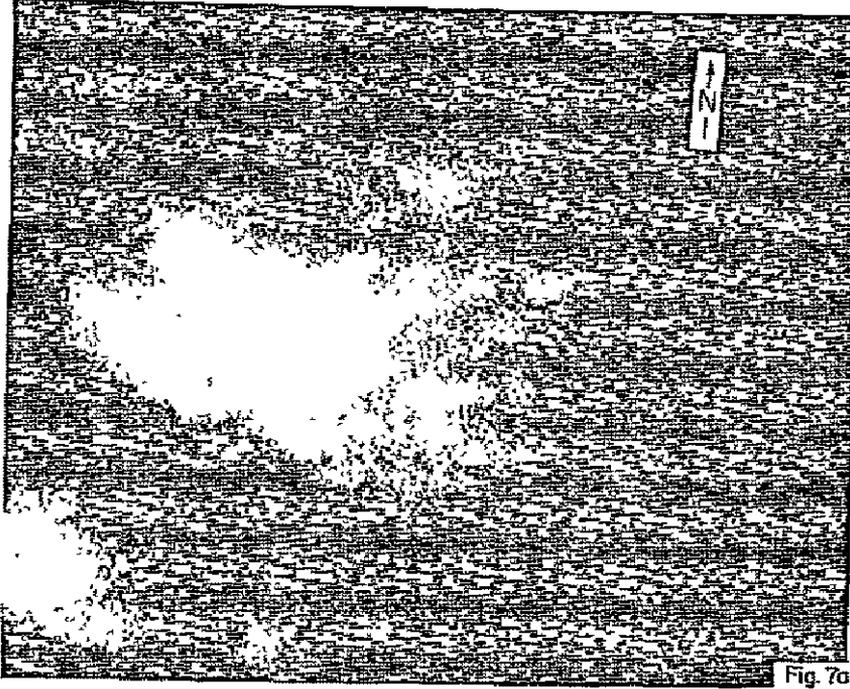


Fig. 7a. Enlargement of a section of large scale cloud field at around 10° S on the southern ITCZ at 2101 GMT on April 19, 1967.

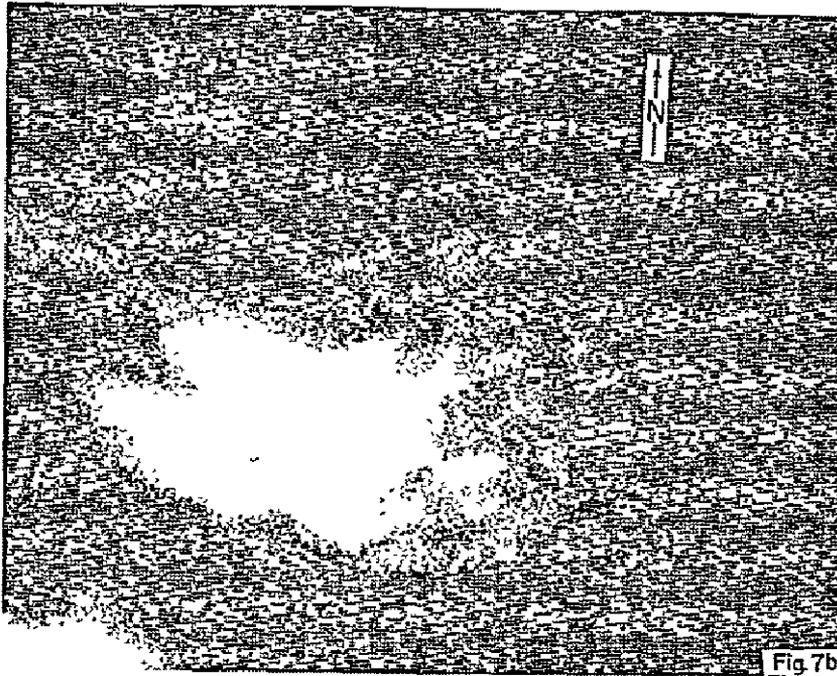


Fig. 7b. Enlargement of a section of large scale cloud field at around 10° S on the southern ITCZ at 2124 GMT on April 19, 1967.

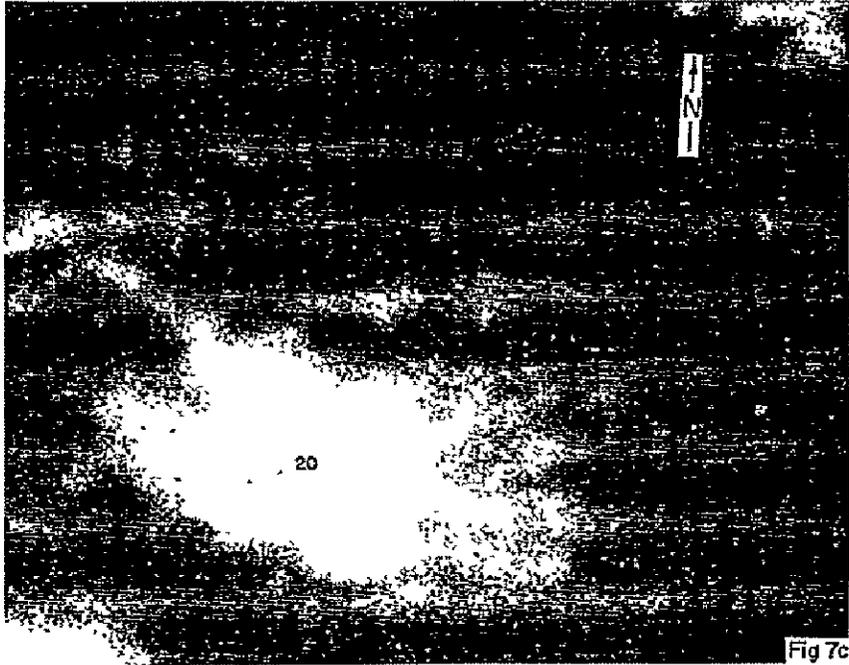


Fig. 7c. Enlargement of a section of large scale cloud field at around 10°S on the southern ITCZ at 2148 GMT on April 19, 1967.

of higher air density where a larger part of the incident energy is lost in the process of atmospheric absorption and scattering.

In general, over the tropical oceans, penetrative convection is associated with cirrus plumes at its tops. These cirrus plumes, when viewed by a geosynchronous satellite, represent anvil cloud material formed out of a thunderstorm complex and exhibit a pronounced brightness. The regions of tallest clouds with maximum updraft in them are the regions of maximum brightness in the cloud complex. The zones of highest vertical motion in a cloud ensemble, in turn, relate the regions of maximum liquid water concentration (Byers and Braham, 1949), for it is in these zones that maximum amounts of water vapor pass into the condensed phase. The shallow convection, on the other hand, exhibit a poor brightness gradient. The key, however, is the growth of the cirrus anvil or plume. It absolutely identifies deep convection areas.

The insets in Figure 5 are examples of a number of isolated convective ensembles. It is clearly seen that the regions of active convection are brighter than the surrounding cloud cover. The cloud masses in these systems are associated with moderate penetrative convection in the ITCZ and as such their sizes are noticeably smaller than those developing in an intense penetrative convection regime. Visual comparison of the cloud brightness in a, b

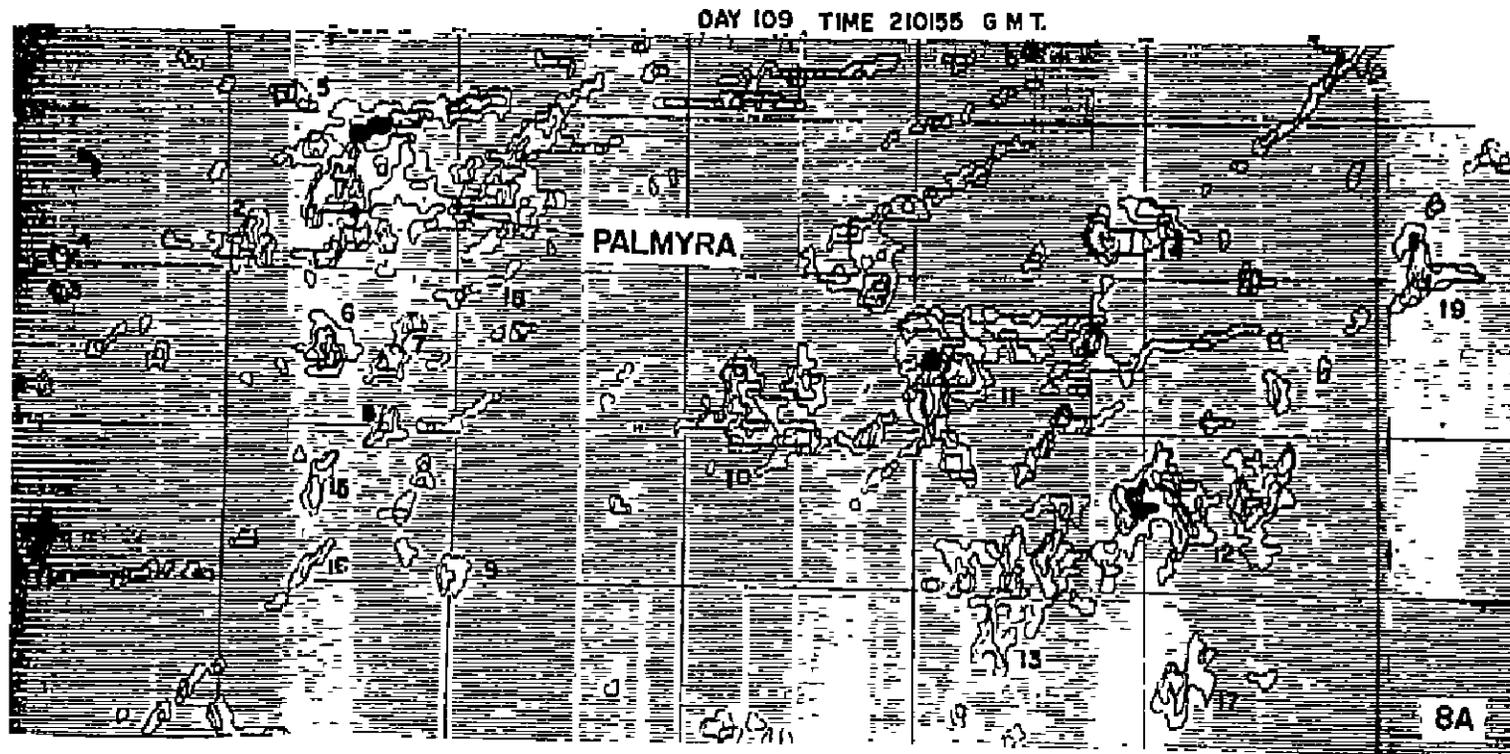


Fig. 8a. ATS-I digital display around Palmyra at 2101 GMT on April, 1967.

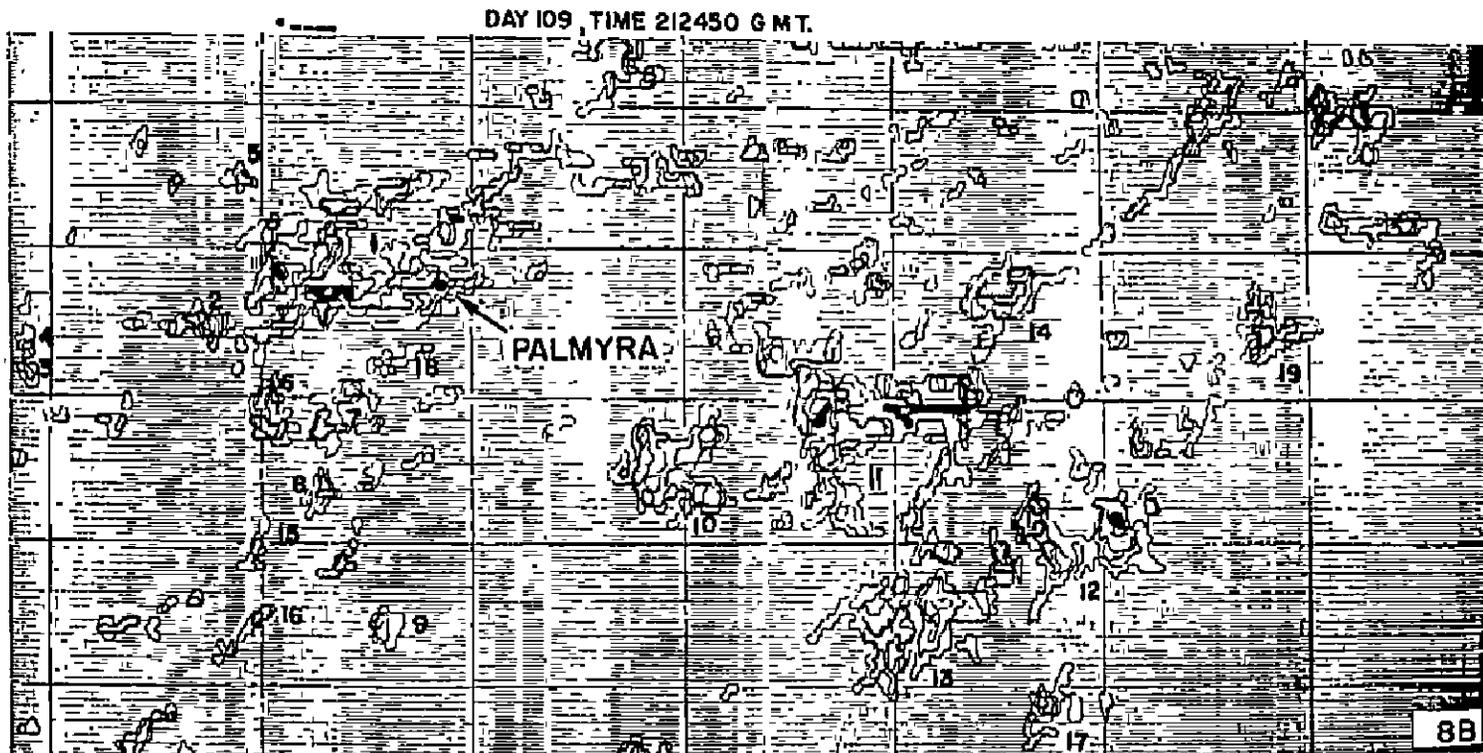


Fig. 8b. ATS-I digital display around Palmyra at 2124 GMT on April 19, 1967.

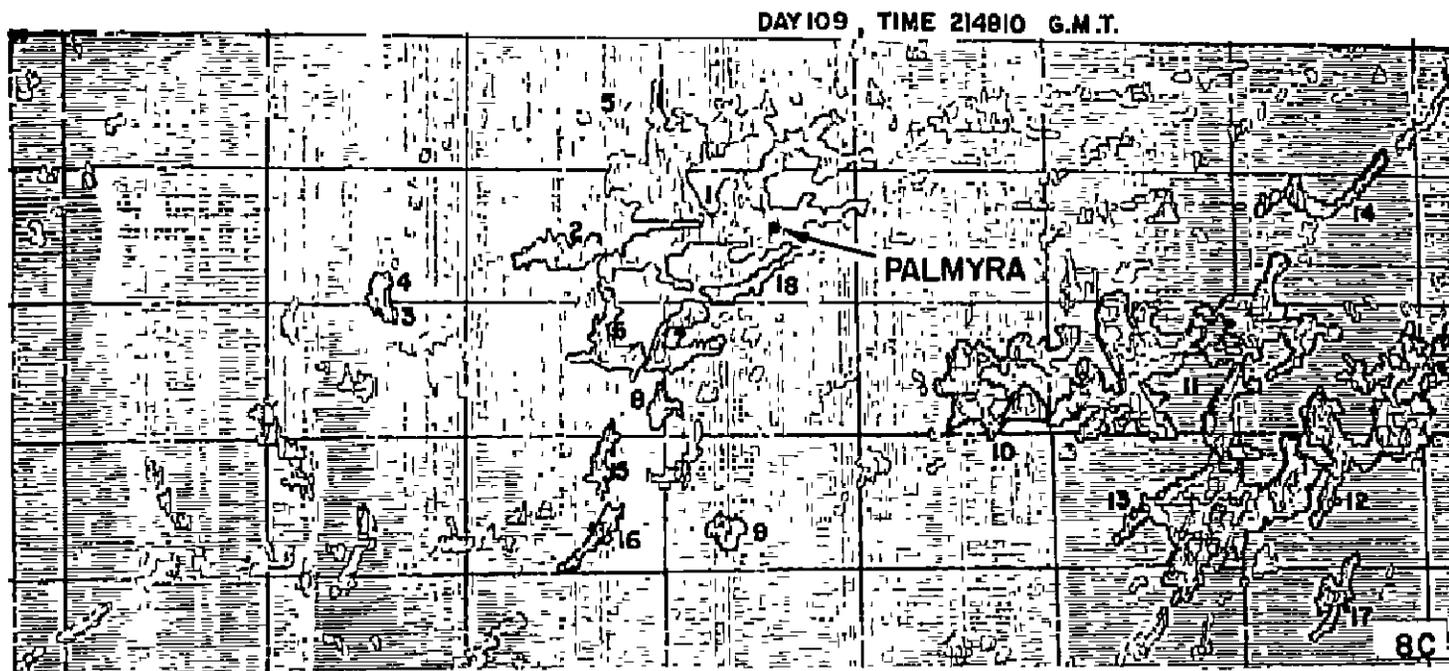


Fig. 8c. ATS-I digital display around Palmyra at 2148 GMT on April 19, 1967.

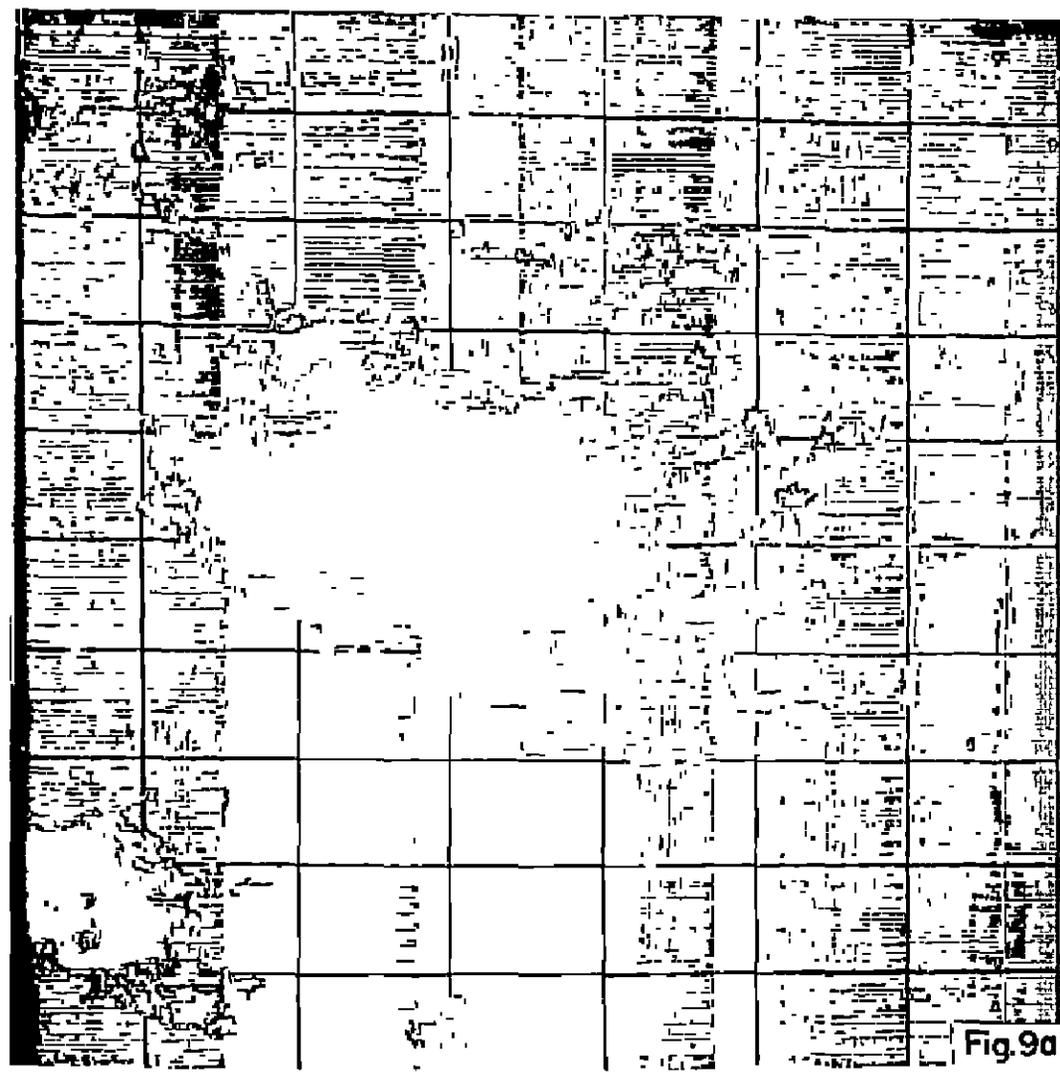


Fig. 9a. ATS-I digital display of the large scale cloud field shown in Fig. 7a.

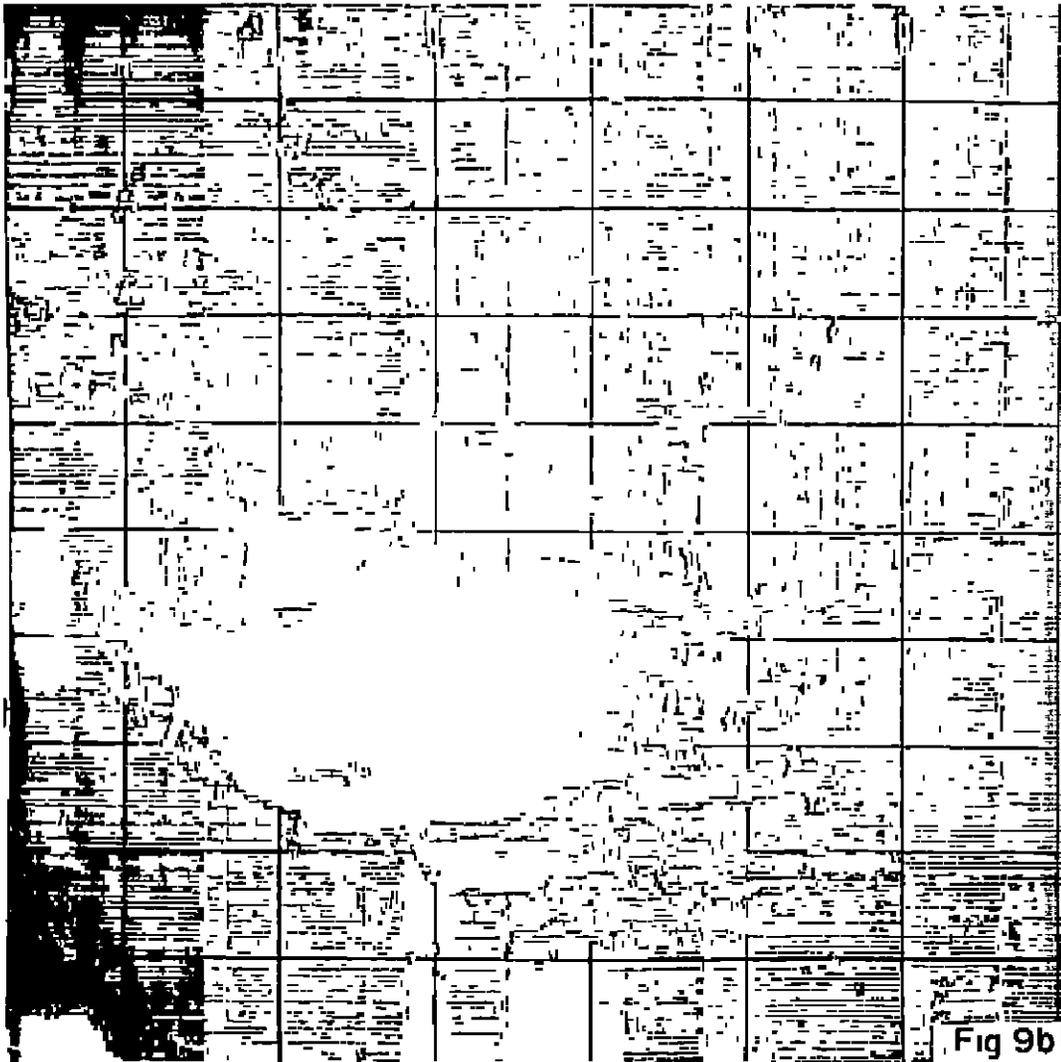


Fig. 9b. ATS-I digital display of the large scale cloud field shown in Fig. 7b.

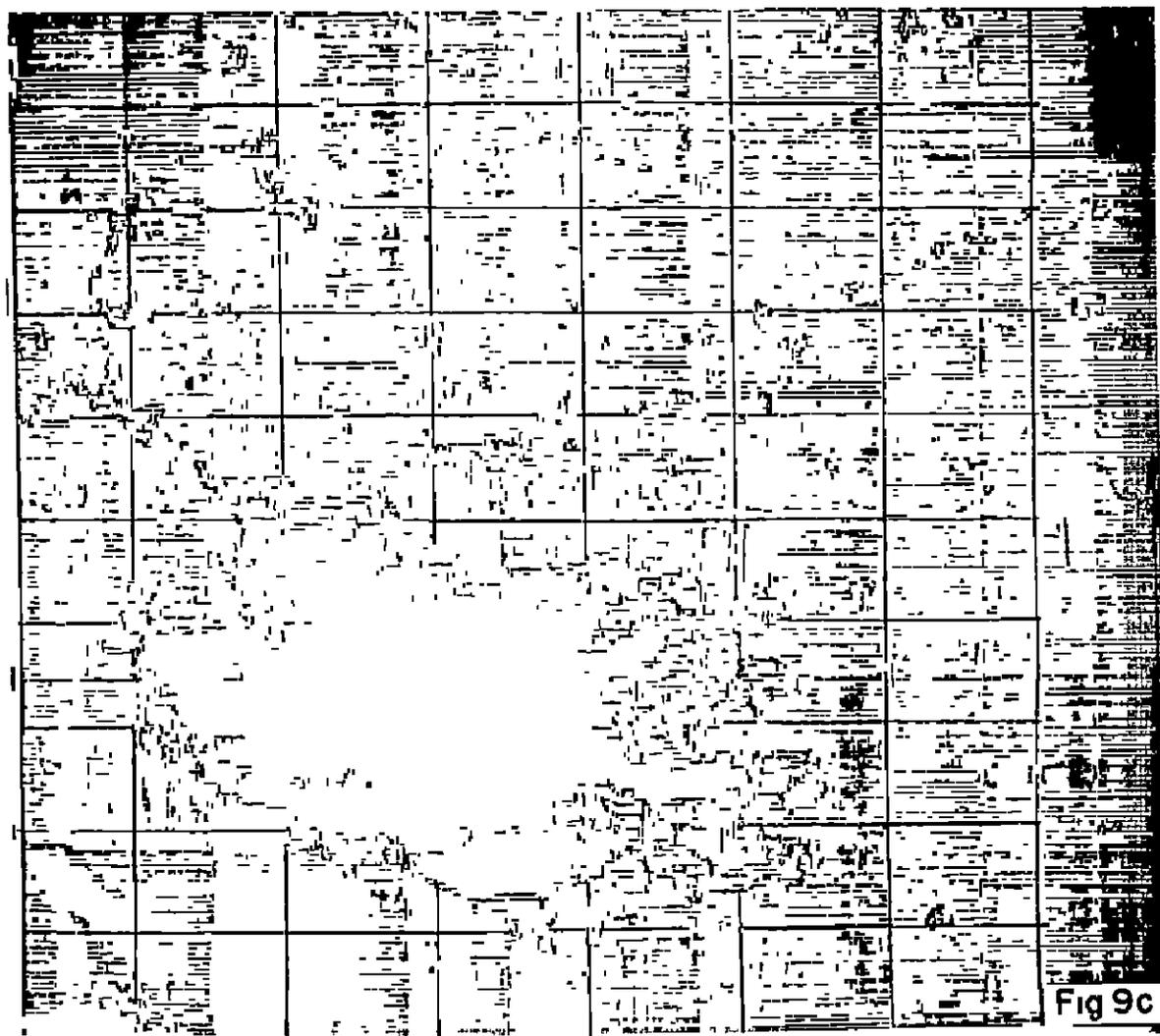


Fig. 9c. ATS-I digital display of the large scale cloud field shown in Fig. 7c.

and c reveals a change in the brightness value with time. This change obviously implies a complex feature associated with the various phases of development of each individual convective cell related to the life cycle of a convective ensemble. The apparent decrease in the brightness of a cloud complex might result from the decay phase of the complex. Some of the convective cells (e. g., cloud #4 in Fig. 7) which were bright in a are less bright in b, becoming brighter in c. These features are typical of a convective complex.

Figures 6a, b, c and 8a, b, c are examples of large scale convection regimes actively associated with the Intertropical Convergence Zone. The strong brightness indicative of the cirrus shield at the top of the convection complex is distinctly noticeable against the low cloud background. In this picture, it is easily possible to outline the cirrus shield but difficult to isolate the zones of active convection in the complex for the brightness gradient is not detected well by a human eye. However, it is possible to isolate them on the ATS-I digital displays.

Once the satellite photographs are properly interpreted for the cirrus shield, it is possible to track the brightest cloud outlines on the two successive satellite pictures. For dynamic reasons, one would expect that the size of these shields is a measure of the potential intensity of convection in a cloud complex since the brightness value is directly related to the upward mass flux of water vapor through the penetrative towers.

Thus, we are able to identify convective regimes of (1) weak to moderate intensity, typical in a mesoscale system, and (2) moderate to strong intensity in large scale convective cloud systems on the ATS-I cloud photographs.

1.6 Comparison with Other Models

In the previous section, we stated that the brightness gradient of a convective cloud plays an important role in deciding the character of convection associated with it. In this section we will use the Line Island data (1967) as the basis for a computation of heat and mass flux using our model and show (1) that our convection model is able to isolate vigorous and moderate to weak convection regime on the ATS-I satellite photographs, (2) that our values of mass and energy flux are consistent with ground-based estimate of Braham (1952) and Brown (1967), (3) that a cloud complex is made of many individual cells, and (4) that this measurement scheme is applicable to storms of moderate to high intensity.

One of the important purposes of the Line Island experiments (1967) was to correlate surface, ship, aircraft and satellite observations. As a "cloud truth" we will discuss the observations in and around the vicinity of Palmyra ($5^{\circ}53'N$, $162^{\circ}05'W$) where we have surface radiosonde, radar and aircraft observations besides ATS-I satellite cloud pictures.

1.6.1 Thermodynamics of the Atmosphere over Palmyra on April 19, 1967

Figure 10 shows the mean upper air temperature and moisture sounding plotted in the Skew-T diagram. The mean sounding is prepared from five soundings taken at 00, 06, 12, 1725 and 2000 GMT. The mean surface temperature is as high as 28.7°C and the specific humidity at the surface is around 19 g/kg becoming less than 0.1 g/kg at 200 mb. The LCL is at 950 mb and the CCL is hardly 10-15 mb higher.

From the radar photographs presented in Figs. 11a, b, c it appears that there were more than 35 precipitation echoes present in the 100 miles radius around Palmyra (5°53'N, 162°05'W). The position of Palmyra on the digital display could not be fixed precisely for want of navigational data relevant for meso-scale analysis. However, an approximate position has been indicated by a black dot on the digital displays. It is interesting to note from the radar photographs that there were as many as 20 precipitation echoes in the cloud complex 1 (see Figs. 8a, b and c) in the first period. These precipitation echoes closely correspond to the positions of the brightest zones on the satellite pictures. The exact superposition is not attempted for there is a difference in the scale.

Nevertheless, it is clear that where conventional observations show the convective activity, the satellite photos show it also. Moreover, there is a good correspondence between the location of radar echoes and the bright spots in the cloud photos. The aircraft observations also indicated scattered cumulus congestus developing with tops above 6-7 km in the west sector of Palmyra close to the period of observations.

Figure 12 presents the mean Q-profile over Palmyra on April 19, 1967. Here the total heat content Q is defined as the sum of enthalpy (C_pT), geopotential energy (gz) and the latent heat (Lq). This diagram reflects on the redistribution along the vertical of heat gained by the atmosphere from the ocean. The Q profile indicates two minima of magnitude, 78.3 cal/gm at 800 mb and 77.9 cal/gm at 600 mb. The level of first minima corresponds approximately to the level of trade inversion around 5°N. The level of second minima may be the highest level through which vertical mixing in the lower troposphere extends by eddy diffusion. The mechanism of eddy diffusion explains the heat and moisture transport from the boundary layer to the middle troposphere, but it fails to account for the counter gradient heat transfer in the upper troposphere if the boundary layer is considered to be the primary heat source. Riehl and Malkus (1958) attempted to resolve this anomaly by assuming the existence of selective buoyancy mechanisms in the tropical convection regime. Cirrus outflow, which hardly occurs below 300 mb in the tropics is evidence of this mechanism. Quantitative estimates of the heat transfer to the upper troposphere by this mechanism is the basic contribution of this paper.

A. T. A. CHART 20 SKEW T DIAGRAM

PRINTED BY THE U.S.A. LAKE METEOR CO., CINCINNATI OHIO U.S.A.

TEMPERATURE IN DEGREES FAHRENHEIT AND CELSIUS

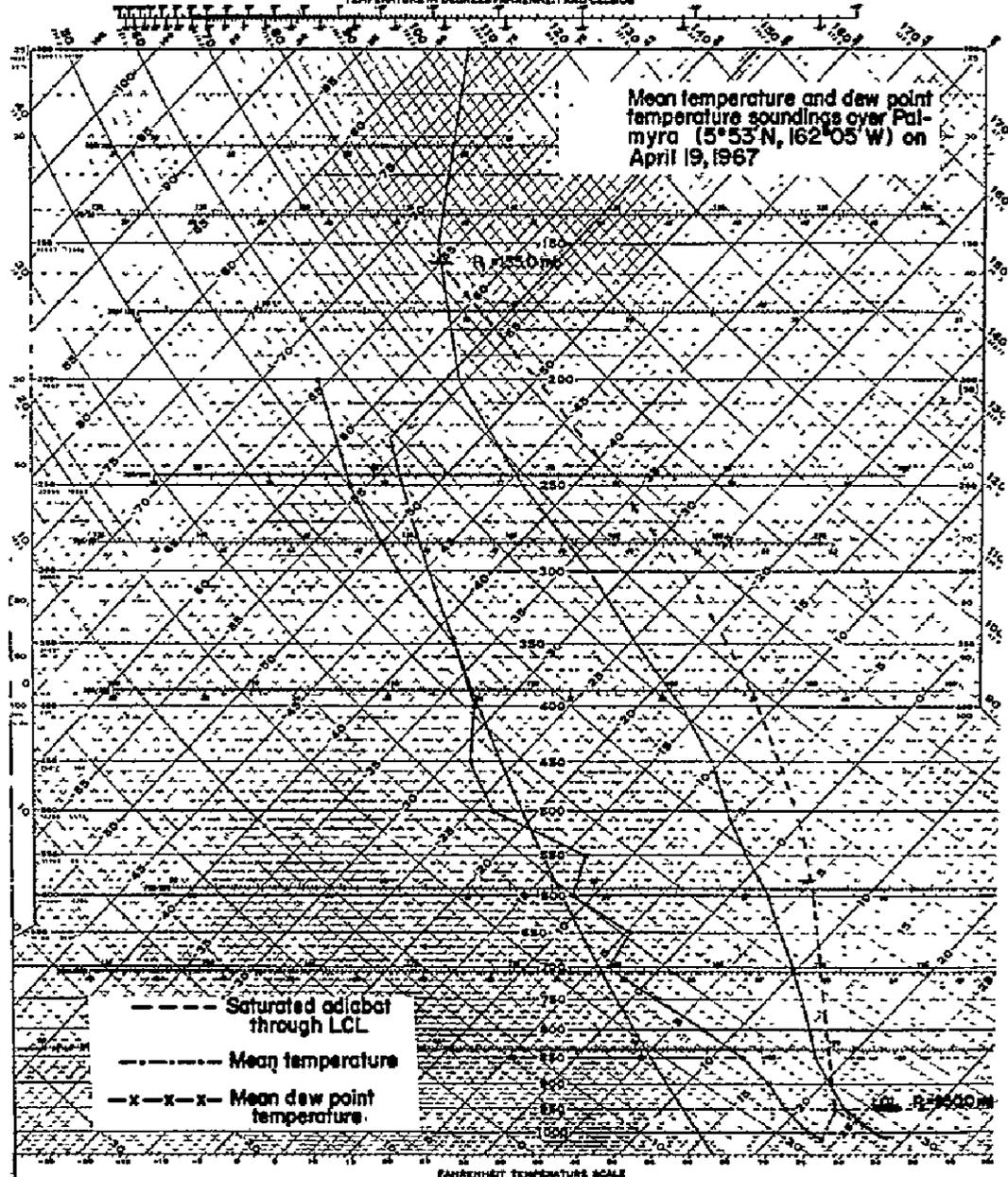


Fig. 10. Mean temperature and dew point temperature soundings over Palmyra (5° 53' N, 162° 05' W) on April 19, 1967.

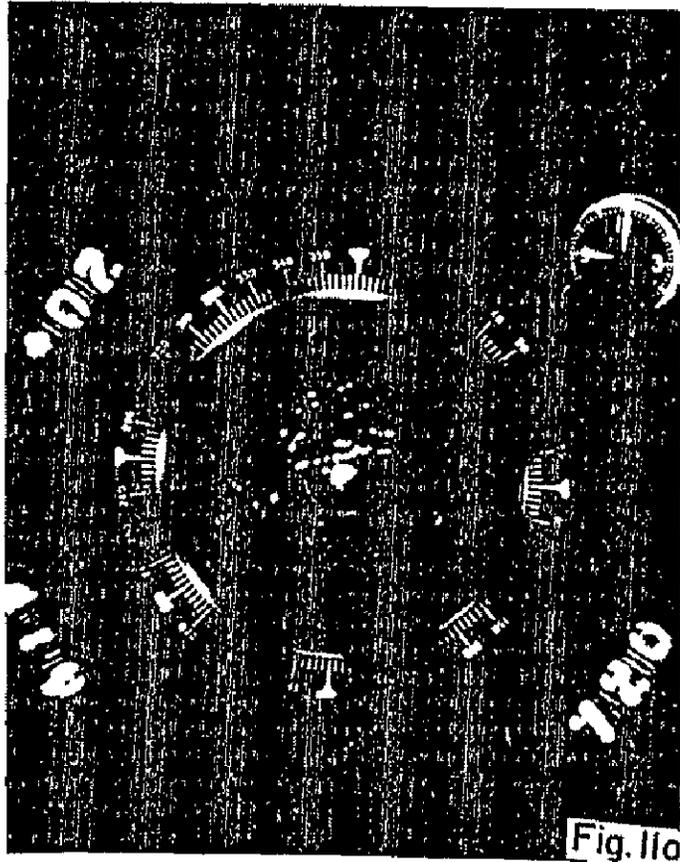


Fig. 11a. Radar echoes around Palmyra at 2100 GMT on April 19, 1967.

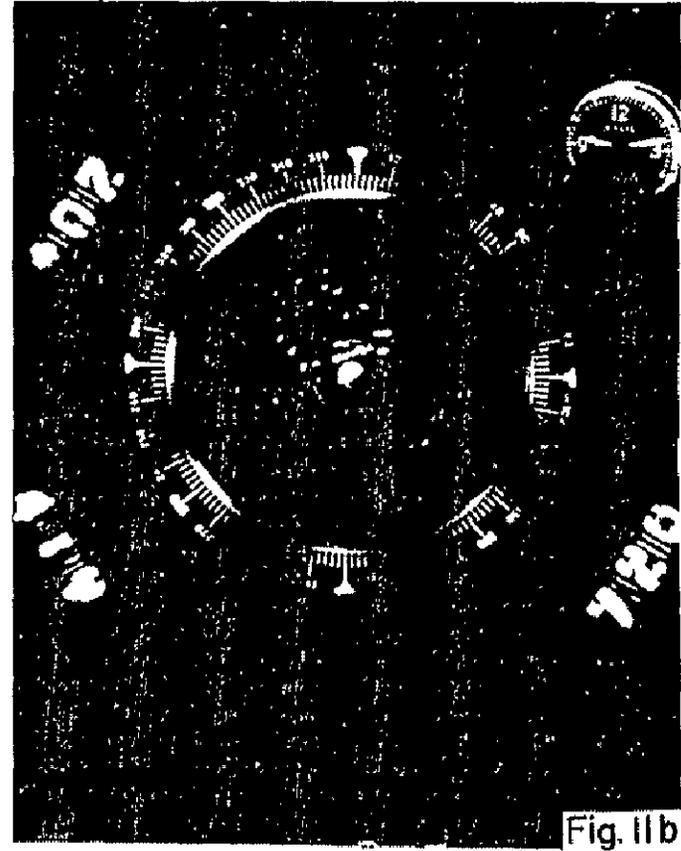


Fig. 11b. Radar echoes around Palmyra at 2114 GMT on April 19, 1967.

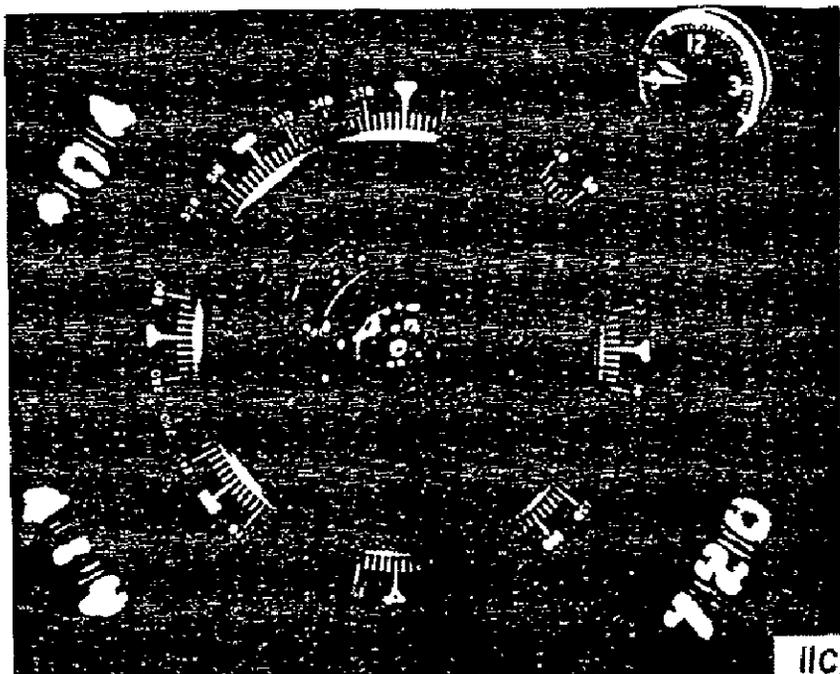


Fig. 11c. Radar echoes around Palmyra at 2145 GMT on April 19, 1967.

1.6.2 Mass Flux Data

Cloud physicists have observed that the growth rate of the convective cloud systems varies considerably from cumulus humilis to cumulonimbi. Brown (1967) has recently compiled all published data on convective cloud growth rates (Anderson, 1960; Brown, 1966), and developed a power function to estimate the growth rate and available energy. The range of mass flux from ground-based measurements has been shown to vary from 10^5 kg/sec for cumulus humilis to $10^7 - 10^8$ kg/sec and above for cumulonimbus clouds.

Figure 13 is a plot of our computed values of mass upward fluxes as a function of cloud area change as viewed by a geosynchronous satellite. All measurements are given in Appendix 1. In Fig. 13, the cloud numbers are indicated against each data point. The numbers in the parentheses refer to the order of the time interval. It is clearly seen from this diagram and Table I and II in Appendix 1 that the mass upward flux, which is really a measure of the convective cloud growth rate, linearly increases with the cloud size and that the values obtained from the ATS-I photographs are entirely consistent with the ground-based estimates. In the light of Brown's (1967) convective cloud mass flux classifications, the clouds indicating flux values greater than $10^7 - 10^8$ kg/sec should be considered as cumulonimbi. The mass flux range 10^6 to 10^8 kg/sec will refer to the category of cumulus congesti. Braham (1952)

gives the mass influx associated with a moderate thunderstorm as 5.1×10^7 kg/sec between the levels 5,000 ft to 10,000 ft.

In the cloud complex 1 in Figures 8 a, b, c, the computed mass flux values are 3.57×10^8 kg/sec and 1.51×10^8 kg/sec during the first and second periods, respectively. These magnitudes are well within the cumulonimbus scale. The decrease of mass flux in the second period is in close agreement with the decrease of precipitation echoes on the latter radar photographs (see Fig. 11c).

Thus, it is clear that the cirrus outflow as seen by a geosynchronous satellite in a moderate to intense convection zone results from many individual cells embedded in the convective complex and that the model evaluates the average magnitudes of the fluxes.

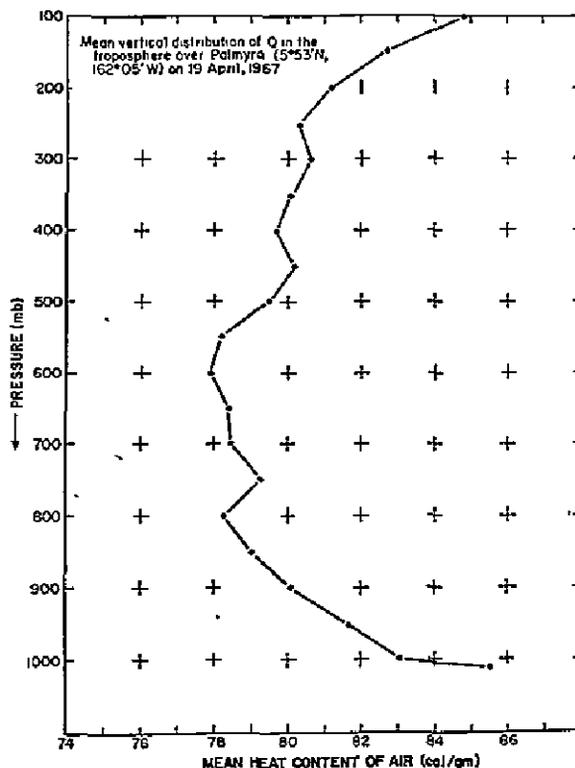


Fig. 12. Mean vertical distribution of heat content of air over Palmyra on April 19, 1967.

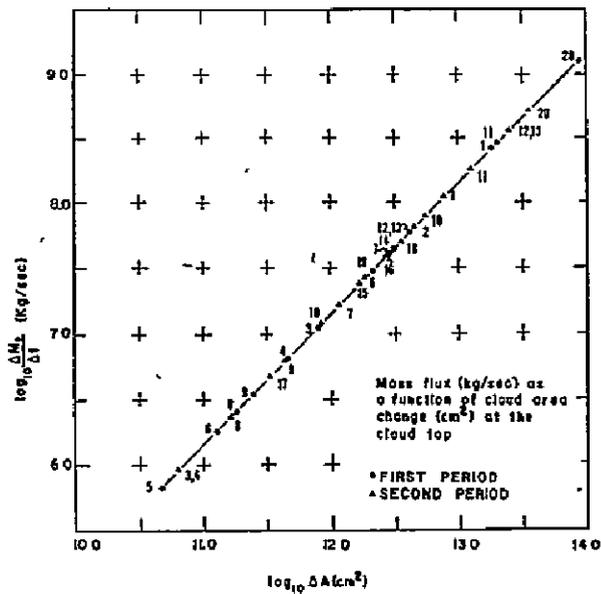


Fig. 13. Mass flux (kg/sec) as a function of cloud area change (cm²) at the cloud top.

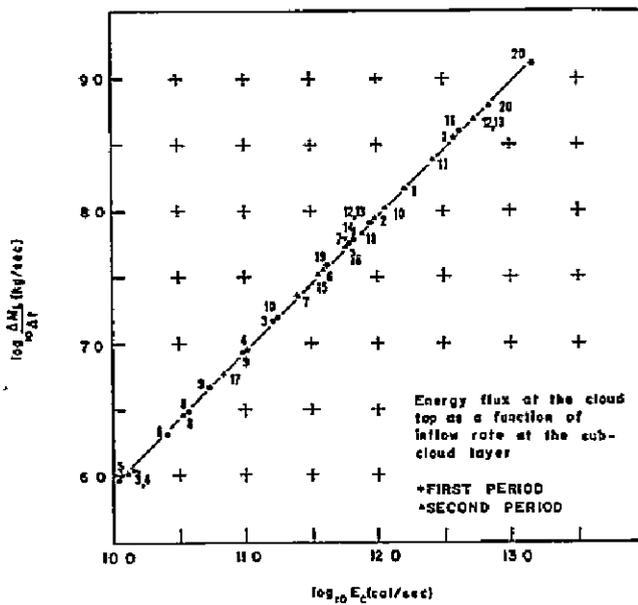


Fig. 14. Energy flux at the cloud top as a function of inflow rate at the sub-cloud layer.

1.6.3 Energy Flux Data and the Cloud Mass

Figure 14 presents the computed energy fluxes as a function of mass influx at the subcloud layer. Figure 15 gives the convective cloud energy (ergs) as a function of cloud mass. The data on the cloud mass are evaluated using Brown's (1967) empirical equation

$$E_c = 1.4 \times 10^9 M_a^{1.14}$$

where E_c is energy (ergs) and M_a is the cloud mass in kg. The estimated cloud mass should refer to the clouds that extend to the upper troposphere and represent the average mass of the clouds participating in the cirrus outflow. According to Brown's (1967) convective cloud mass scale (see Fig. 16), a single cumulonimbus should have cloud mass greater than 10^{22} ergs. Braham also quotes the same order of magnitude for the total work done in an average thunderstorm against the environment. In the cloud complex 1, where towering cumulus congesti were reported on the west section of Palmyra by an aircraft earlier to the period of observations, the computed energy values are 22.3×10^{22} (ergs) and 9.48×10^{22} (ergs) in the first and second periods, respectively.

Referring back to Tables 1 through 4 in Appendix 1, one finds that the cloud number 20 has the largest contribution towards the upward mass and heat transport to the tropical upper troposphere. The computed magnitudes of heat flux for this cloud are 0.67×10^{13} cal/sec and 4.3×10^{13} cal/sec during the first and second periods, respectively. These magnitudes are nearly two orders higher than the evaluated for mesoscale complexes. This implies that the energy transport in the large scale cloud processes in the tropics is significantly large compared to isolated mesoscale convective complexes.

It is, therefore, clear from the data presented that one can, indeed, estimate the average magnitudes of convective fluxes of mass and heat in the tropics from the ATS-I cloud photographs based on the model and classify the scales of convection accordingly.

The technique outlined in this part of the thesis may as well be employed for the estimation of horizontal momentum exchange through deep convections in the tropics. This problem, however, has not been treated in this research.

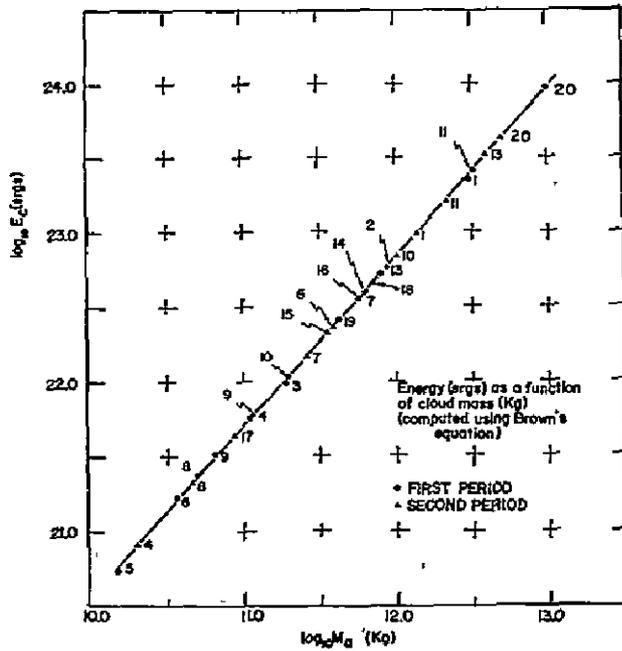


Fig. 15. Energy (ergs) as a function of cloud mass (kg) (computed using Brown's equation).

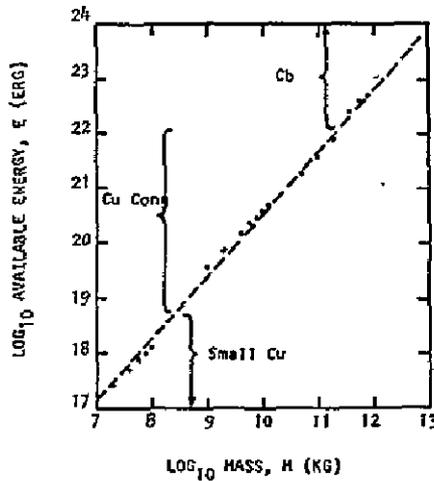


Fig. 16. Available energy as a function of mass of convective cloud. Dashed line is best-squares fit to data points. (Brown, 1967)

2. A SATELLITE ESTIMATE OF LARGE SCALE CONVECTIVE TRANSPORT OF MASS AND ENERGY

2.1 Introduction

In the first part of the thesis we have developed a tool for probing into the energetics of a convective complex in the tropics as seen by a geosynchronous satellite. This part will deal with the application of the technique for the evaluation of average transport of mass and heat through large scale convective systems in the mid-Pacific.

The tropical oceans have long been recognized as a heat and moisture source of the atmosphere. Most of the mass and heat transport to the tropical upper troposphere first pointed out by Hadley is accomplished through deep convection. Riehl and Malkus (1958) estimated the heat budget parameters for the equatorial trough zone based on the radiation estimates of London (1957). Their estimates demanded an upward mass transport of 18×10^{13} g/sec through undiluted hot towers to the tropical upper troposphere in order to compensate for the radiative losses. For the entire 10° latitude belt around the mean position of the ITCZ, Riehl and Malkus (1958) have shown that only 1500-5000 active giant clouds are needed at a time to maintain the heat balance of the equatorial trough zones.

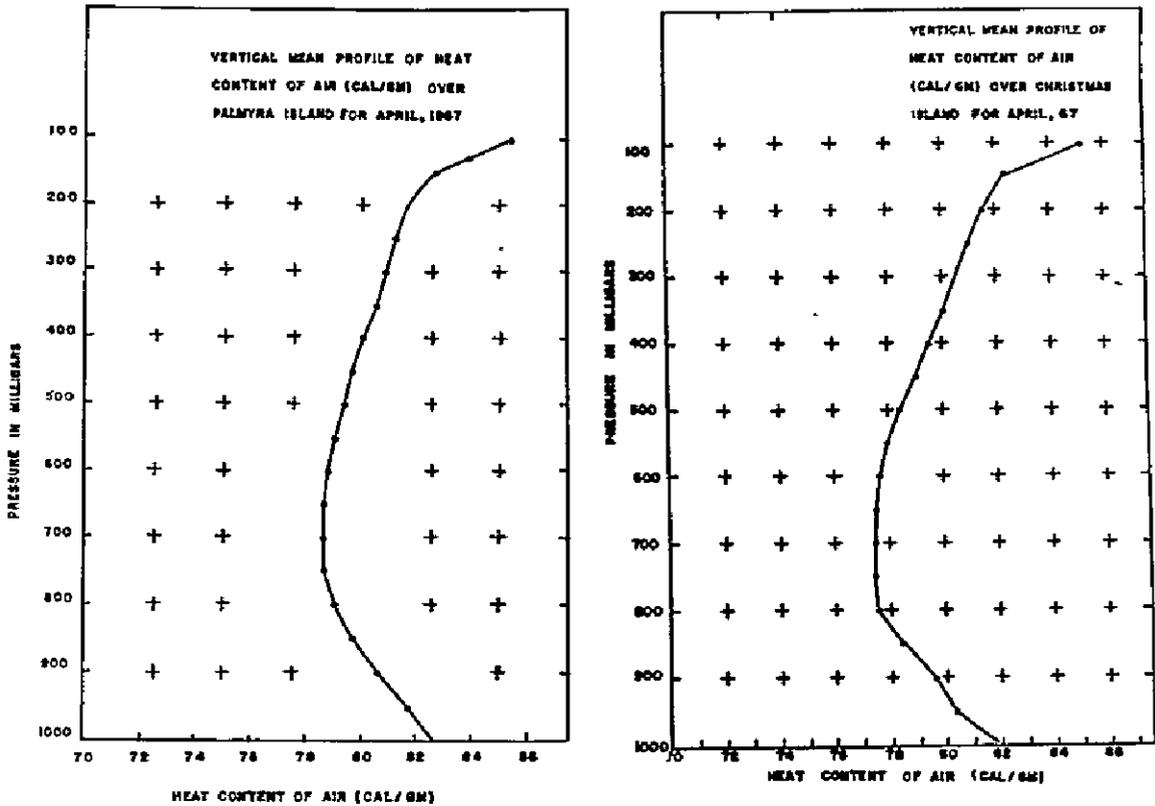
Recently Vonder Haar (1968), using more than forty months of radiometric measurements from the first generation meteorological satellites, has shown that in the low latitudes the net radiation is about 40 percent higher than previously estimated by London (1957). In the light of Vonder Haar's (1968) estimate of radiation budget parameters one should expect much greater transport of heat from the tropics.

The "hot tower" concept of Riehl and Malkus (1958) leads to a broader understanding of the mean heat sources and sinks in the atmosphere but their analysis did not look into the short and long period fluctuations of these transports that might result from various large scale dynamical causes. Also, their study did not answer the question as to how these meso- to sub-synoptic scale convective cloud processes are coupled to the large scale heat transfer poleward from the tropics.

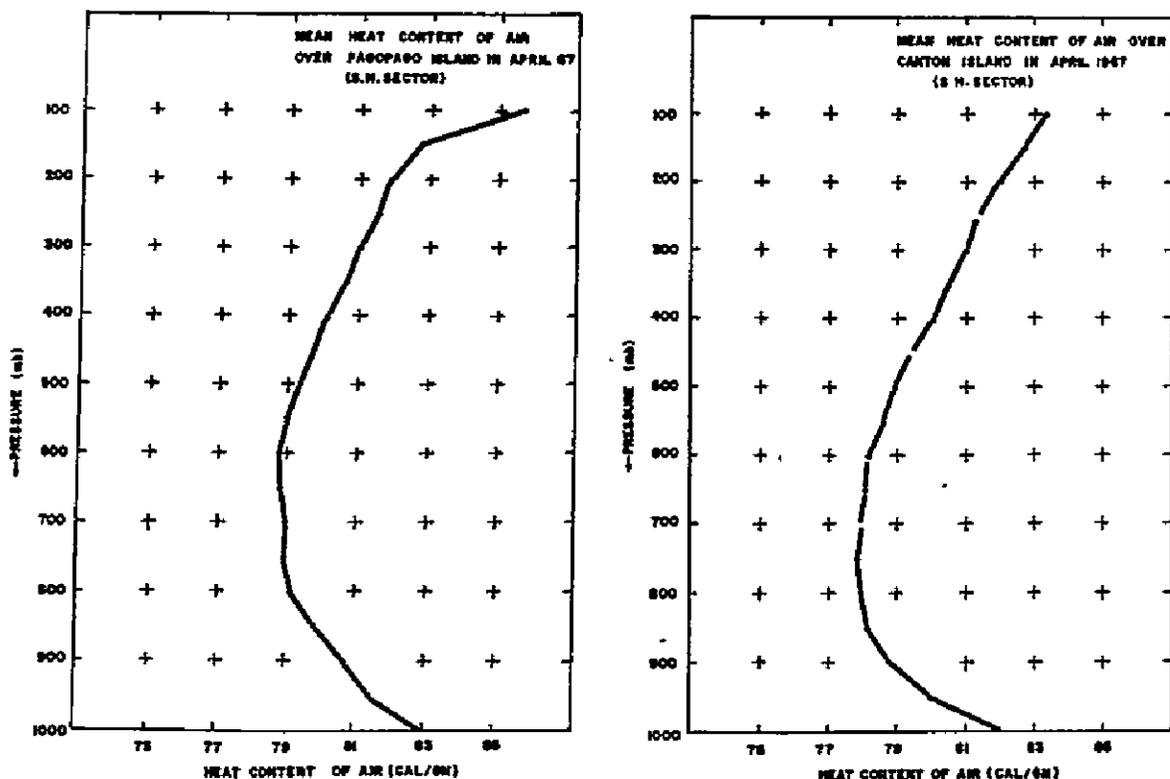
In this part is an attempt (1) to estimate, from the ATS-I satellite cloud photographs, the sensible heat and latent heat transfer through the undiluted "hot towers" in the tropical convergence zone, (2) to investigate as to how these computed large scale fluxes vary with time, and (3) to what extent these fluxes are comparable with others' estimates.

2.2 Implications on the "Hot Tower" Concept

Figures 17a and 17b present the mean vertical cross-sections of total heat content of air ($Q = C_p T + gz + Lq$) over Palmyra ($5^{\circ}52'N$, $162^{\circ}06'W$), Christmas Island ($2^{\circ}N$, $157^{\circ}23'W$), and Figs. 18a and 18b over Canton Island ($2^{\circ}46'S$, $171^{\circ}43'W$), and Pago Pago ($14^{\circ}S$, $171^{\circ}W$) for April, 1967. The minimum Q -values in these cross-sections is higher at Palmyra Island than over Christmas Island. The interesting part of these diagrams is the increase of Q -values above 500 mb. The existence of higher Q -values in the tropical upper troposphere cannot be explained by a simple mass circulation of the Hadley cell as pointed out by Riehl and Malkus (1958). The mechanism of diffusive advection through small scale convections may account for the transport of heat and mass to the level of minimum Q , the average top of the moist layer, but it fails to explain the counter-gradient heat and mass transport to the tropical upper troposphere. Embedded in the tropical convergence zones are cumulonimbi clouds whose central cores are protected from mixing with the environment by the large cross-sections of the clouds. We have shown in Part I



Figs. 17a, b. Vertical cross-sections of Q over Palmyra and Christmas Island in the N. H. S.



Figs. 18a,b. Vertical cross-sections of Q over Palmyra, Christmas Island in the N. H. S. and Pago Pago, Canton Island in the S. H. S.

that the cirrus outflow from these undiluted towers in the large scale cloud field are easily identified on the ATS-I satellite photographs.

2.3 Grid Selected in this Study

The grid used in this computation scheme covers the area between longitude 120°W to 180°W and latitude 15°N to 15°S over the mid-Pacific ocean. The section 120°W to 180°W , 0°N to 15°N will, henceforth, be referred to as the Northern Hemisphere (N.H.) sector, while the region 120°W to 180°W , 0° to 15°S becomes the Southern Hemisphere (S.H.) sector. In this grid box, it is assumed that there is no significant change in the cloudiness from the lateral advection into the grid box over a period of 2-3 hours and that the large scale wind fields do not change appreciably during the period of observations.

The observation period is planned for two hours on either side of local noon at the sub-satellite point, because during this period of small solar zenith angle, the cloud brightness is nearly independent of solar zenith angle

within the sector. In the period of 4 hours we can get as many as eleven cloud photographs, provided the satellite was turned on to obtain continuous coverage.

2.4 ATS-I Satellite Data Coverage

Although the ATS-I satellite is capable of a continuous watch over the Pacific province, because of other scientific experiments on the spacecraft this is rarely done. During the Line Islands Experiment (LIE) (March, April, 1967), however, an almost continuous watch was maintained for a comparatively longer period.

Table I presents the data covering during the observation period (i. e. , two hours on both sides of local noon at the sub-satellite point) each day for April, 1967. It is seen from this table that the data coverage is far from satisfactory during the period from April 1 through April 12, except on April 2 and April 4, 1967.

To compute the vertical heat transfer, we need at least two satellite cloud photographs. In this computation scheme we have used all available information with not less than three photographs in the observation period, except on April 12, 1967. For want of sufficient samples the daily means of the average magnitudes may not be very representative for certain days, but when considered over a period of one month it is presumed that these individual contributions will be adequate.

The study of short period fluctuations of the large scale cloud parameters will be confined to the period April 16 through April 23, when there was no single interruption in the satellite watch.

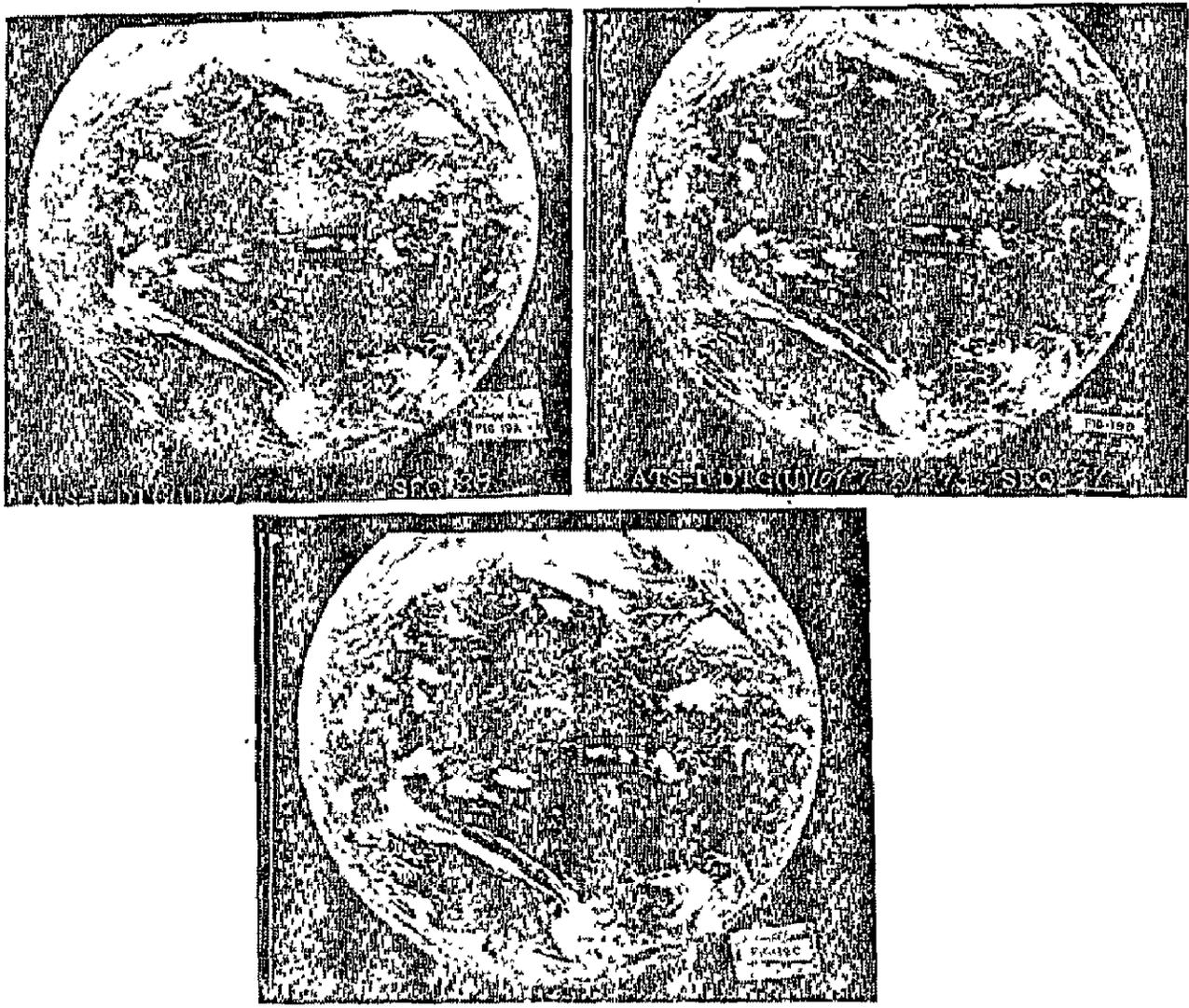
2.5 Technique of Analysis

The bordered regions in Fig. 19a, b, c show a section of a large scale cloud field associated with the Southern ITCZ. These photographs are presented to show the significant difference in the brightness values between the mesoscale and large scale cloud complexes in the tropics. Figures 20a, b, c are the enlargements of the bordered section in Figs. 19a, b, c. The large scale cloud fields are well defined on the ATS-I satellite cloud photographs. These are the areas of ascending air connected with cloud cluster disturbances in the tropics. Martin (1968) showed that "cloud clusters" are an identifiable entity of the tropics and the term has since been adopted in a Joint Organizing Committee working group on tropical studies. The areas covered by these large scale cloud fields associated with the synoptic disturbances at one time are small—hardly 1/10th of the area of the tropics. From the ATS-I satellite cloud photographs, one can easily visualize that the upward transport of heat

TABLE I

Number of Cloud Photographs Used in This Scheme

<u>Date</u> <u>April, 1967</u>	<u>Day No.</u>	<u>No. of Cloud</u> <u>Photographs</u> <u>Used</u>	<u>Remarks</u>
1	91	3	
2	92	8	
3	93	0	Data inadequate
4	94	10	
5	95	3	
6	96	4	
7	97	3	
8	98	4	
9	99	3	
10	100	5	
11	101	0	Data inadequate
12	102	2	
13	103	12	
14	104	11	
15	105	7	
16	106	12	
17	107	12	
18	108	12	
19	109	11	
20	110	11	
21	111	12	
22	112	12	
23	113	13	
24	114	7	
25	115	10	
26	116	11	
27	117	3	
28	118	8	
29	119	4	
30	120	<u>4</u>	
Total		217	



Figs. 19a, b, c. A section of the large scale cloud field in the southern ITCZ.

A SECTION OF A LARGE SCALE CLOUD FIELD
ON THE SOUTHERN HEMISPHERE ITC Z

DAY 109, APRIL 19, 1967



(a)

TIME 21060 Z

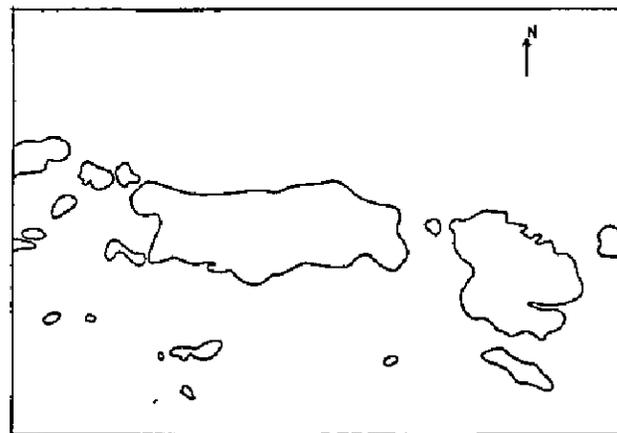


(b)

TIME 212455 Z

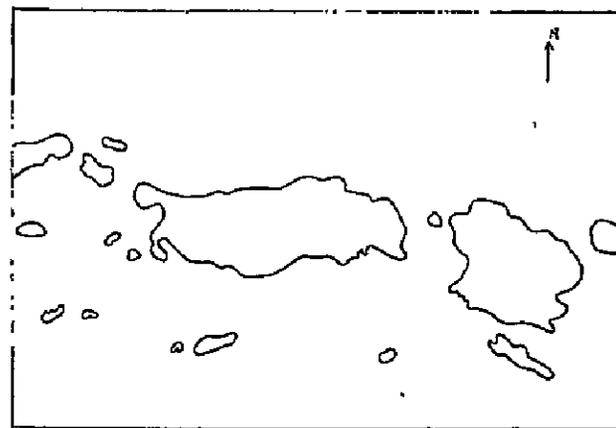
Fig. 20. Enlarged section of the bordered section in Figs. 19a, b, c.

CLOUD ANALYSIS



(a)

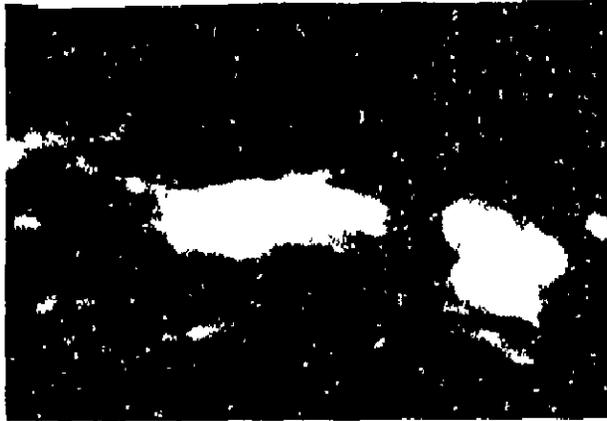
TIME 21060 Z



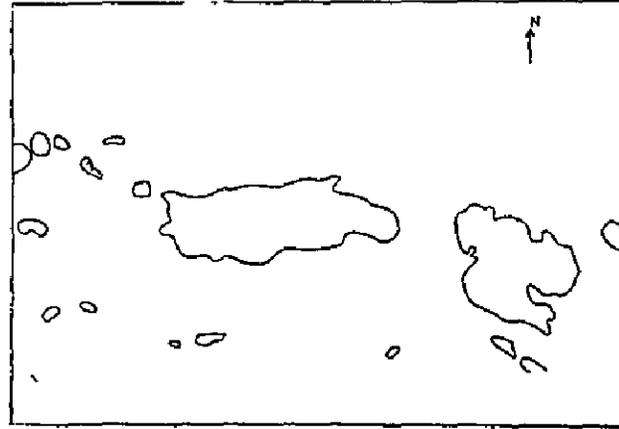
(b)

TIME 212455 Z

Fig. 21. Outlines of the enlarged cloud in Figs. 19a, b, c.

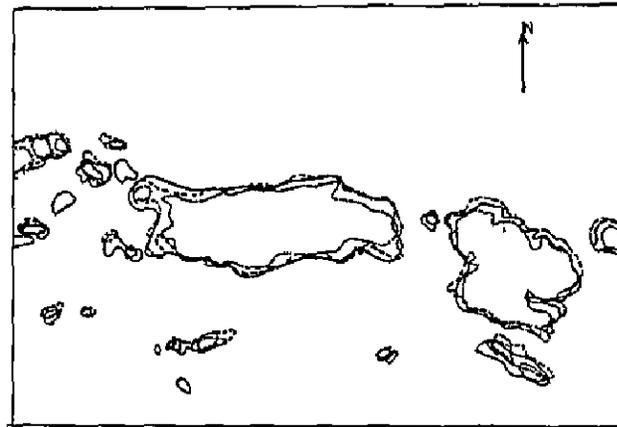


TIME 21400 Z
Fig. 20(c)



TIME 21810 Z
Fig. 21(c)

CLOUD AREA CHANGE WITH TIME BASED ON
THE ABOVE PHOTOGRAPHS



— 21050 Z
- - - 21245 Z
— 21400 Z

Fig. 22

and moisture must be accomplished in these zones over the tropical oceans.

As already mentioned in Part I, the intense brightness of this large scale cloudiness is pronounced at the tops of the deep clouds. Cirrus outflow is common in the zones of intense convection and the size and particularly the growth of the cirrus shield is an indication of its severity. The regions of brightest cloud cover on successive ATS-I satellite cloud photographs can be easily outlined and tracked.

Figures 21a, b, c are the examples of these outlined made from the ATS-I satellite photographs shown in Figs. 20a, b, c, respectively. Figure 22 is the superposition of the cloud outlines shown in Figs. 21a, b, c. Figure 22 shows significant changes in the cloud area within a period of 44 - 45 minutes.

This technique of heat transfer estimate has been applied to ATS-I satellite cloud photographs over the tropical Pacific. Example photographs are presented in Fig. 23, and details are given in Appendix 2 in the form of tables. The enlarged photographs show some intense but small scale convection cells whose areas are not large enough to be planimetered. We have already shown that the area change of scale convective cells is small enough compared to that of large scale systems to be ignored.

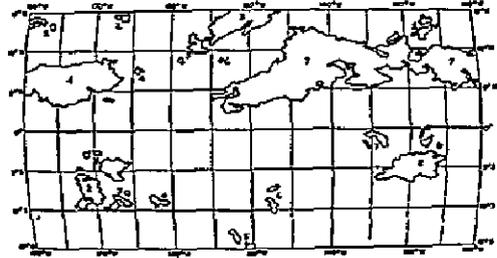
2.6 Vertical Cross-section of Temperature and Moisture in the Grid Box

Our computation scheme requires information on the mean vertical cross-section of temperature and moisture profiles in the grid box shown in Fig. 24. We have only a few meteorological stations taking upper air soundings routinely in the study area, but changes are comparatively slow in the tropics. We have used upper air observations recorded from Palmyra Island ($5^{\circ}52'N$, $162^{\circ}06'W$), and Christmas Island ($2^{\circ}N$, $157^{\circ}23'W$) in the N. H. sector, and Canton Island ($2^{\circ}46'S$, $171^{\circ}43'W$), and Pago Pago ($14^{\circ}S$, $171^{\circ}W$) in the S. H. sector. The mean vertical profiles of $T^{\circ}C$ and $W/g/Kg$ from all of the four stations for April 1967 are presented on a thermodynamic diagram (Skew-T), Fig. 25, while the mean over individual stations in the N. H. and S. H. sectors are given in diagrams (Figs. 26 and 27), respectively. These diagrams indicate that the horizontal variations in temperature at any level are insignificant except in the case of Christmas Island which shows a temperature profile more than $1.0^{\circ}C$ lower than the mean at all levels. This temperature gradient southward may be sufficient to give a mass circulation between these two regions.

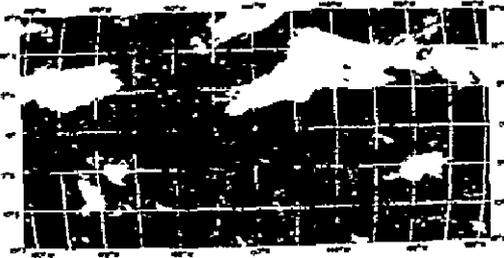
We have taken $17.9 g/Kg$ as the mean subcloud layer mixing ratio for our computation scheme. The mean lifting condensation level (LCL) in the grid box is about 950 mb, and the convective cloud top appears to be at 150 mb. The presence of a large positive instability area from the LCL to the cloud top is conducive to a convection once the parcel is triggered to the level of LCL.



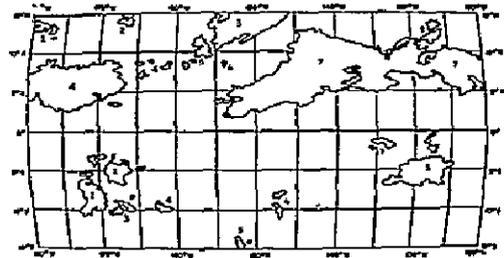
TIME 194900 Z



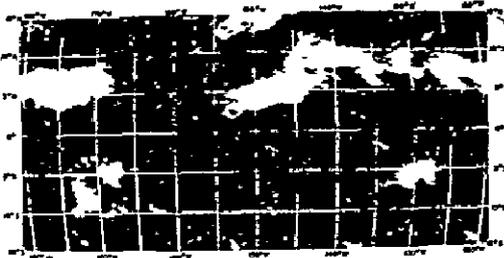
TIME 194900 Z



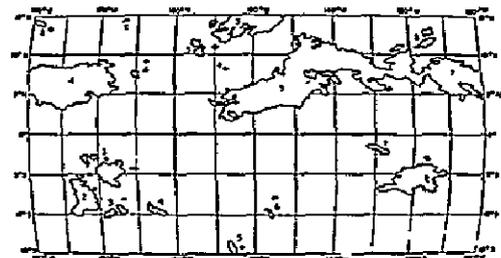
TIME 200500 Z



TIME 200500 Z



TIME 203430 Z



TIME 203430 Z

Fig. 23. A sequence of ATS-I cloud photographs of the day 107 used in this research scheme.

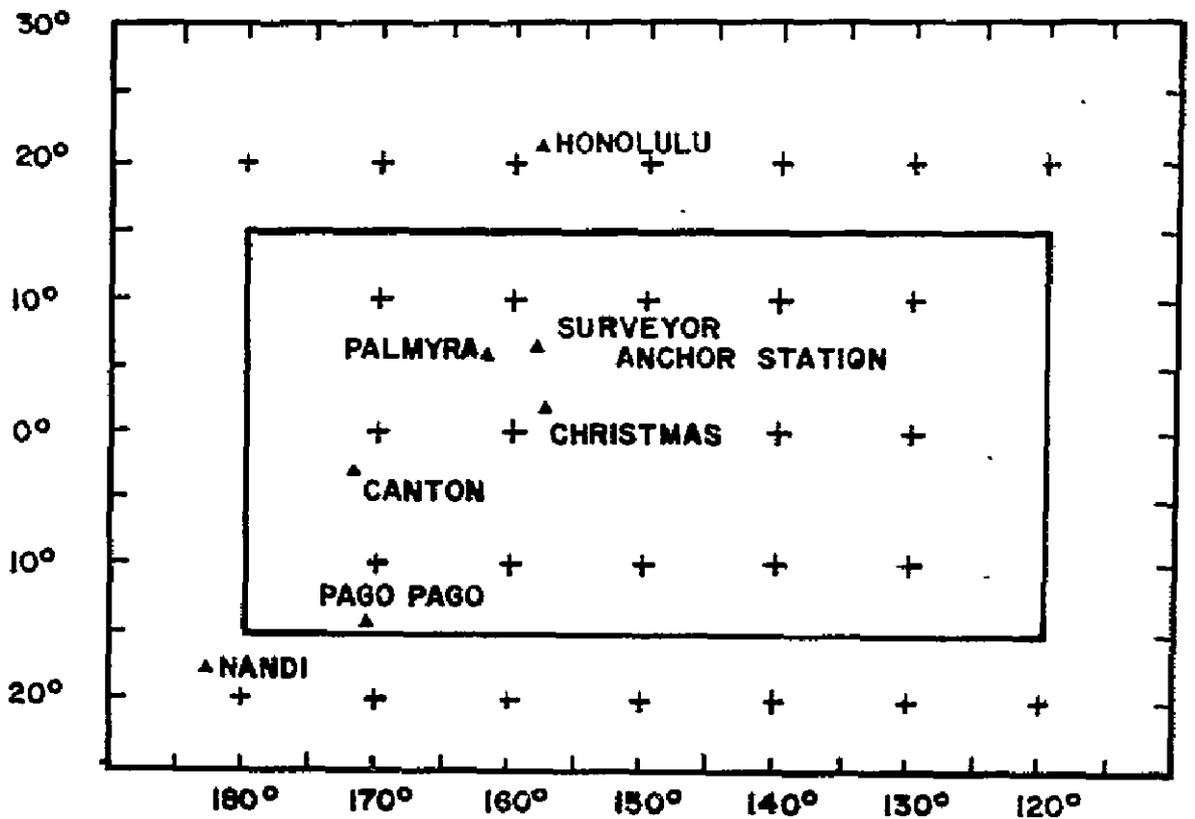


Fig. 24. Location of radiosonde stations in the grid box under investigation.

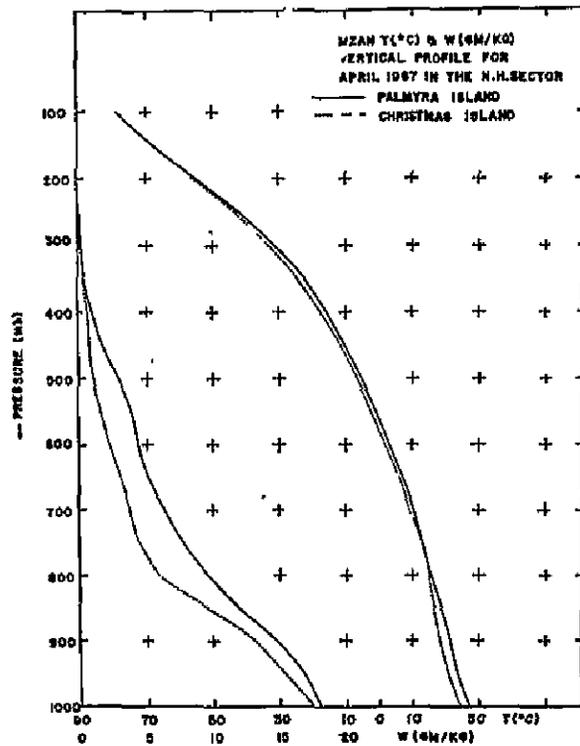


Fig. 26. Vertical cross-sections of temperature °C and mixing ratio (g/kg) over Palmyra and Christmas Island.

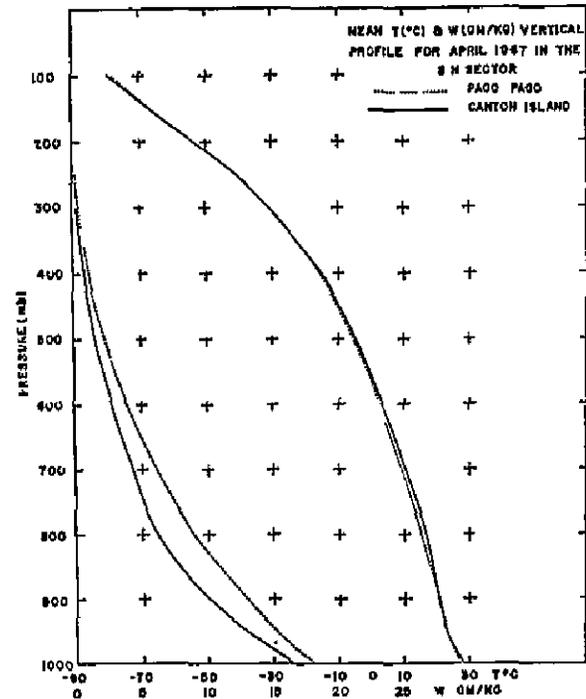


Fig. 27. Vertical cross-sections of temperature °C and mixing ratio (g/kg) over Pago Pago and Canton Island.

2.7 Short Range Fluctuations of the Meteorological Events

2.7.1 Percent-bright Cloud Cover

Figures 28 and 29 present the percent-bright cloud cover in the N. H. S. and S. H. S. as a function of time for the observation period on days #106 through #113. Figures 30 and 31 show their deviations from the mean as a function of time. Off hand one would expect to find a short time variation in the cloud cover for the short time scale of ordinary convections. These diagrams show that such changes do exist but we are unable to find any relationship with diurnal heating or atmospheric tides. All one can say from these observations is that on some days more active convections are present than on other days.

2.7.2 Average Upward Energy Transport

Figures 32a, b present the average upward transport as a function of time in the N. H. S. and S. H. S., respectively. It is expected that the heat flux in the grid box would fluctuate in the same pattern as the cloud tower growth-rate. The growth-rate of a convective complex is reflected in the rate of cirrus shield expansion. These diagrams clearly point out that the convective activity varies during the day but not on the diurnal cycle, and that the large scale upward heat transport is accomplished in the form of surges. These heat surges seem more frequent and larger in amplitude in the N. H. sector than in the S. H. sector of the grid box. In a period of 3-4 hours the heat released during these surges may be as high as 1-2 langley per minute over an extensive area.

2.8 Long Range Fluctuations of Meteorological Events in the N. H. and S. H. Sectors

2.8.1 Cloudiness Cycle

While our studies have shown that during the day the convective activity is more or less random, the picture is very different when we look at the activity extending over a number of days. On this time scale, we find the convective activity well organized.

Figure 33 presents the daily mean percent-bright cloud cover during the period of observations for April, 1967. The percent-bright cloud cover in the N. H. S. is considerably higher all through the month. The peaks in both the sectors are almost always in phase, but their day-to-day variations do not follow the same trend. The phase differences in the two trends may be attributed to the local differences in the meteorological conditions (e. g., temperature, moisture field in the vertical) in the two sectors. The simultaneous appearance of the peaks in the two sectors probably indicates that

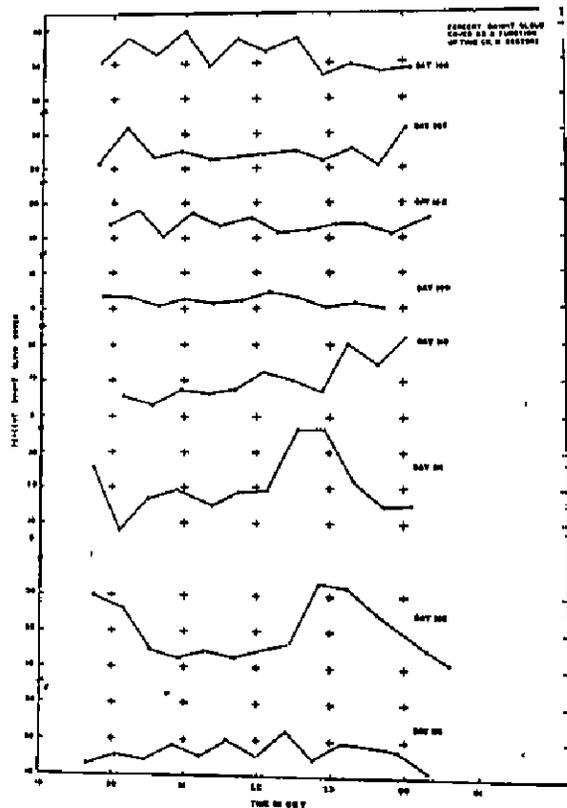


Fig. 28

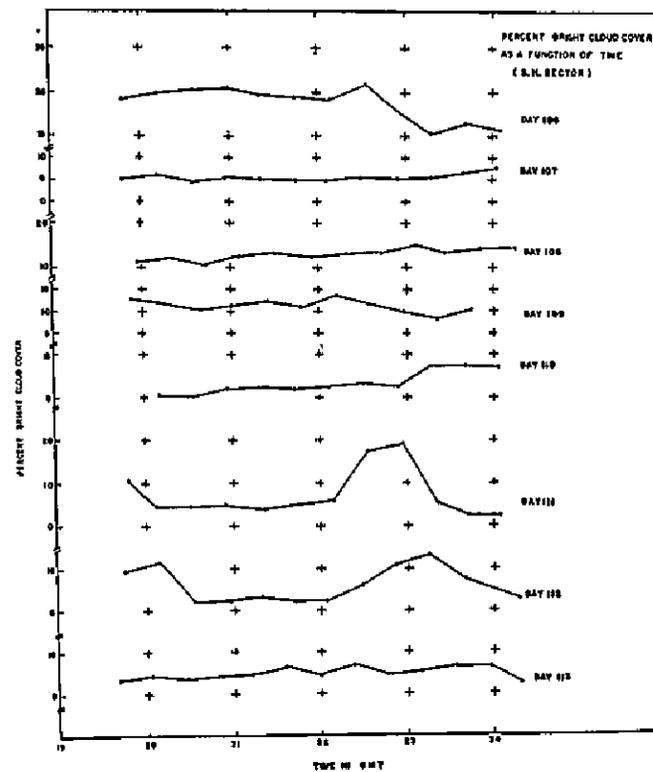


Fig. 29

Percent-bright cloud cover in the N. H. S. and S. H. S. as a function of time.

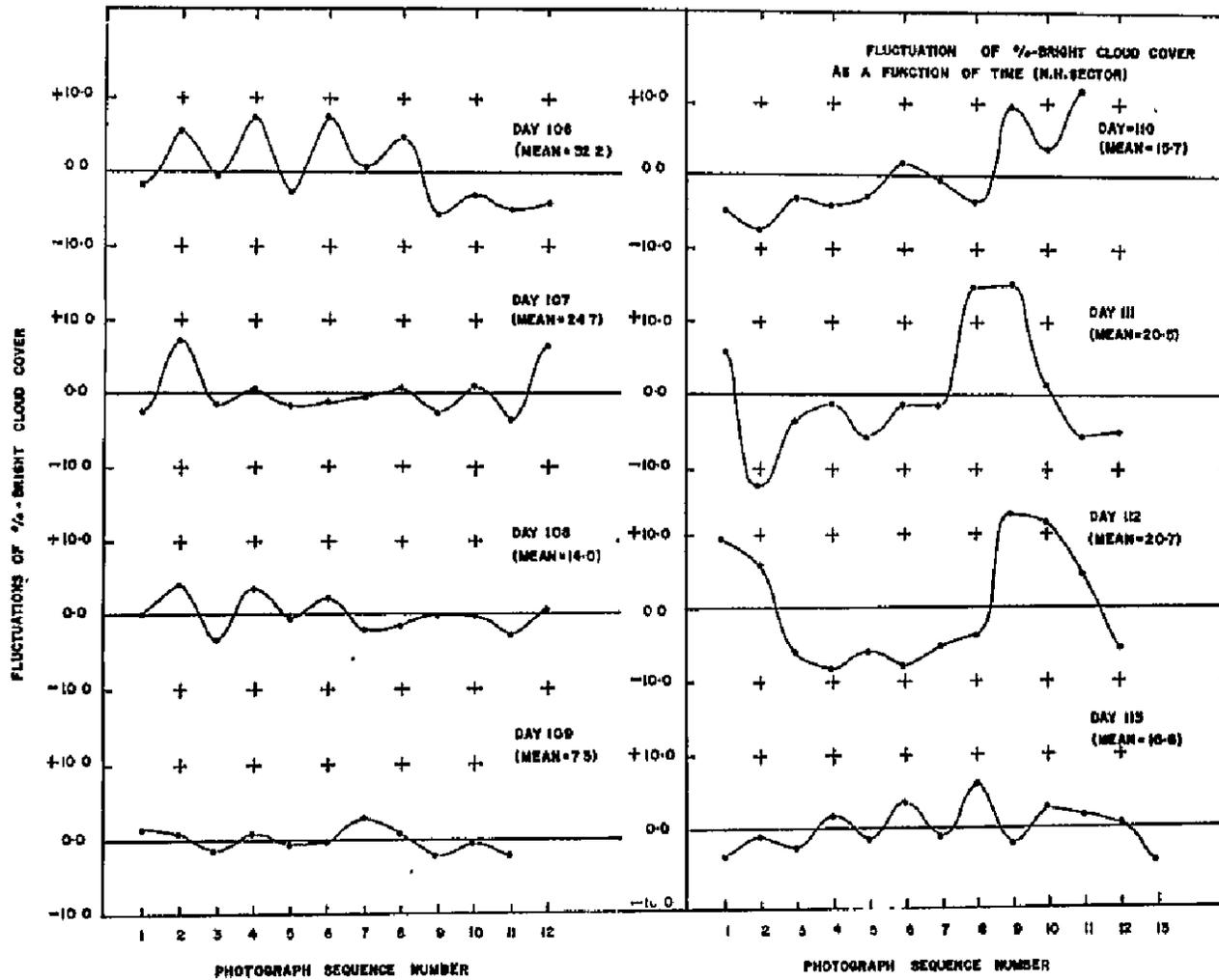


Fig. 30. Fluctuations of bright cloud cover as a function of time (short period).

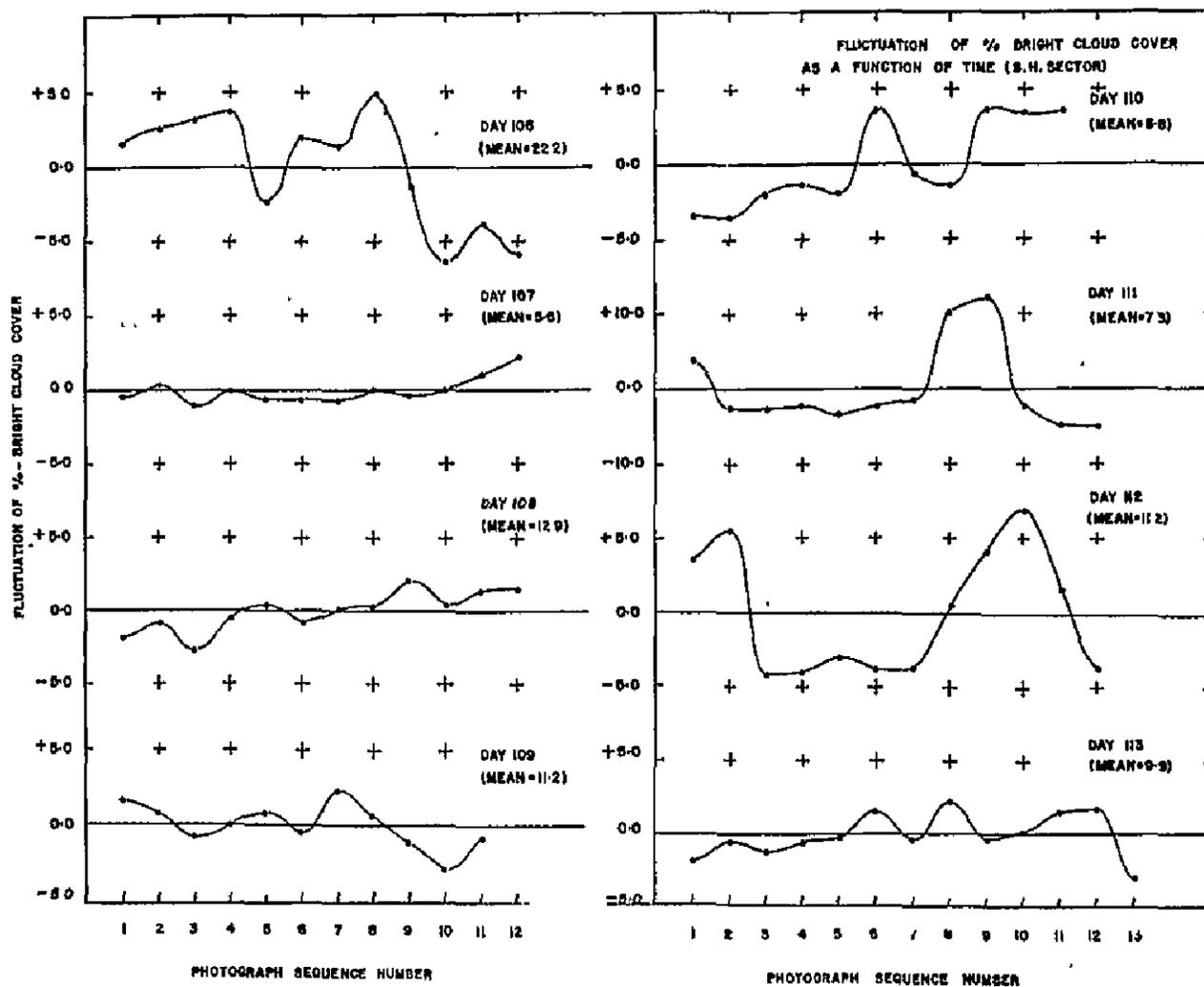
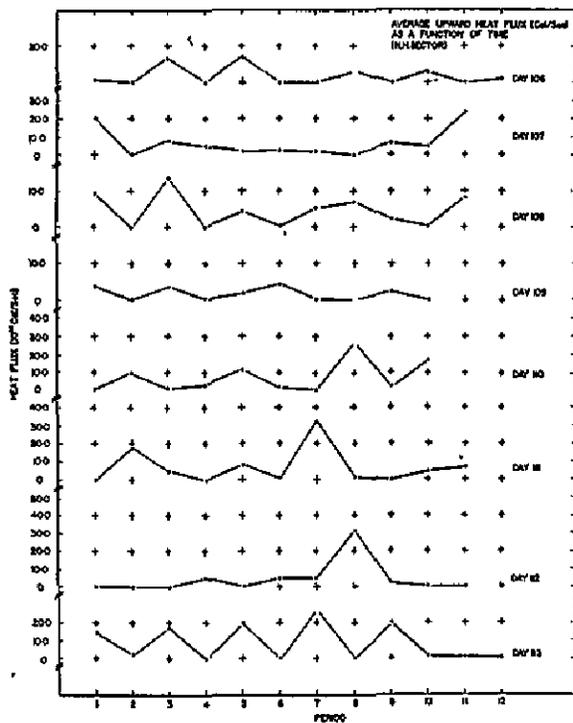
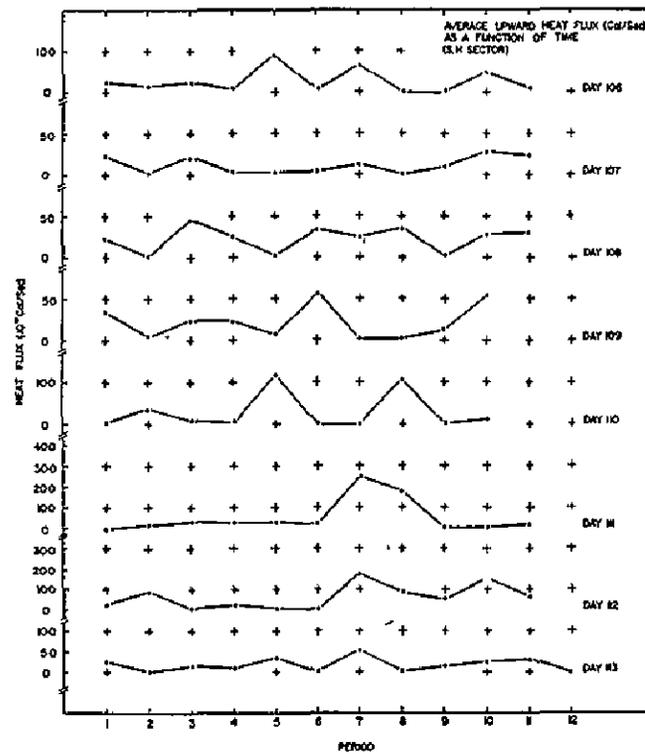


Fig. 31. Fluctuations of bright cloud cover as a function of time (short period).



(a)



(b)

Fig. 32. Average upward energy transport as a function of time (short period).

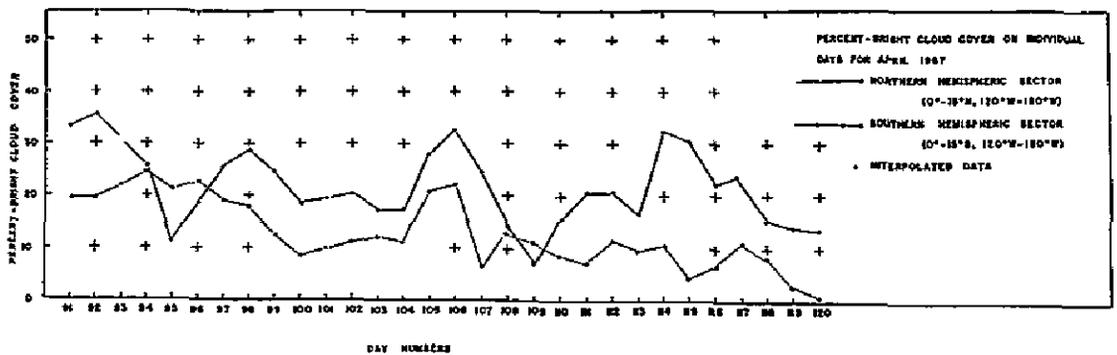


Fig. 33. Daily mean percent-bright cloud cover during April, 1967.

the intensification of large scale cloudiness cycles over the equatorial Pacific is accomplished by the same upper tropospheric disturbance field.

2. 8. 2 Daily Mean of the Average Cloud Divergences

Figure 34 gives the daily mean of the computed average cloud divergences. These values represent the averages of the positive cloud divergences only. The estimate of the percent-bright cloud cover, on the other hand, includes both growing and decaying clouds. The cloud divergences in the layer of out-flow are proportional to the convergence field and their fluctuations reflect on the character of vertical motion field (i. e., weak, strong, etc.) in the zone of intense convection in respective grid sectors. This time cross-section clearly shows that the vertical motion field varies considerably from day-to-day over the tropical mid-Pacific and significantly differs in magnitude in the two sectors.

2. 8. 3 Mass Inflow Rate at the Subcloud Layer and the Convective Heat Transport to the Tropical Upper Troposphere during April, 1967

Table II gives the mean of the average mass influx at the subcloud layer and of the average heat transport to the tropical upper troposphere in both the hemispheric sectors. In the N. H. S., the mass influx in the subcloud layer is about twice that over the S. H. S. The greater inflow rate is an obvious consequence of stronger heat convergence field there at the boundary layer. The estimated mass inflow rate 4.87×10^{13} g/sec at the subcloud layer compares very well with Riehl and Malkus' (1958) data on the lateral influx in the layer 1000-900 millibars. Their estimate pertains to a 10° latitude belt around the globe with the northern ITCZ as the base line and to the month of February.

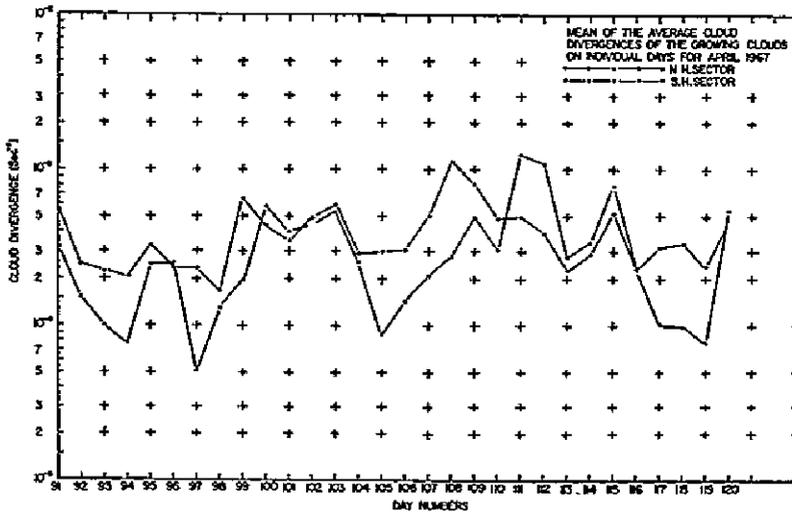


Fig. 34. Daily mean of the average cloud divergences.

TABLE II

Monthly mean for the mass influx rate (10^{13} g/sec) at the subcloud layer and for the convective heat transport (10^{14} cal/sec) to the tropical upper troposphere (April, 1967)

Number of days computed		Mass influx rate		Convective heat transport		Heat flux density (ly/min)	
<u>N. H. S.</u>	<u>S. H. S.</u>	<u>N. H. S.</u>	<u>S. H. S.</u>	<u>N. H. S.</u>	<u>S. H. S.</u>	<u>N. H. S.</u>	<u>S. H. S.</u>
28	28	4.87	2.40	5.24	2.60	0.2838	0.1408

Table II also indicates the computed flux density as 0.2838 ly/min for the N. H. S. and 0.1408 ly/min for the S. H. S. (Sikdar, 1968). The upward transport rate in the N. H. S. is nearly double that in the S. H. S. Considering the area and season these flux densities for April are about 30 to 40 percent higher than those of Riehl and Malkus (1958). House (1965) and Vonder Haar (1968) have shown that the planetary albedo is less in the tropical region and greater at higher latitudes than according to earlier calculations. This demands a greater transport of energy poleward than was thought necessary. In view of their results, the higher values of convective heat transport in this research compared to those of Riehl and Malkus should be expected, but the reader is reminded that Riehl and Malkus calculated the transport for the whole tropical belt

whereas this study pertains to the area 60° longitude \times 30° latitude—about a $1/4$ sample.

Figure 35 shows the daily mean of the average heat transport for April, 1967. The most interesting feature in this diagram is the periodic fluctuation of the convective heat transport rate in both the hemispheric sectors, especially over the N.H. sector. Each data point on this diagram represents the averages of many growing clouds in the respective sectors, and on a number of successive cloud pictures. Although 30 such data points are inadequate for conclusive time series analysis, it is felt that at least some qualitative information on the periodicity of the large scale tropical disturbance field and its interaction with the mesoscale convection systems may be obtained. Figure 36 gives the power spectral estimate of these heat fluxes. It is clear from this diagram that the large scale convective transport pulsates with an approximate periodicity of five days in the N. H. S. while in the S. H. S. the periodicity ranges between 6 to 10 days. These observations invariably imply that the large scale convective systems over the tropical mid-Pacific are dynamically coupled to the equatorial wave type disturbance field (Palmer, 1950, 1951), and that their individual contributions are additive. These observations further imply that the large scale vertical motion in the tropical troposphere results from the integrated effect of an ensemble of rather rapidly ascending cloud groups separated essentially by nondivergent clear spaces.

Thus, it is clear from this study that the daily variations in the convective activity in the tropics are controlled by the large scale motion field. Yanai et al. (1968) indicated from the analysis of power spectra of meridional components of wind in spring over the tropical mid-Pacific that a large spectral density is found at periods 4 to 5 days throughout the upper troposphere where the mean wind is from the west. Our observations are in agreement with Yanai's findings.

In our computation scheme we hypothesized that in the ITCZ there are high intensity "hot towers" that transport moisture to the tropical upper troposphere sufficiently to produce cirrus shields. Our observations supported the hypothesis. The meridional heat exchange between the tropics and extra-tropics is probably accomplished through these cirrus streamers in the tropical upper troposphere which are dynamically coupled to the oceanic boundary layer convergence field. In fact, Suomi and Fujita (1968) presented direct evidence through a time lapse picture from the ATS-III that narrow streams of cirrus originated over tropical cloud systems over the Pacific, flowing north-eastward with a velocity as high as 100 knots.

Further, these large scale cirrus streamers when spread laterally near the tropopause may remain dynamically inactive but radiatively effective to produce strong radiative convergence through the cloud layer source. Over a long time scale these streamers may act as a radiative source in which case our estimates of heat fluxes may be even greater.

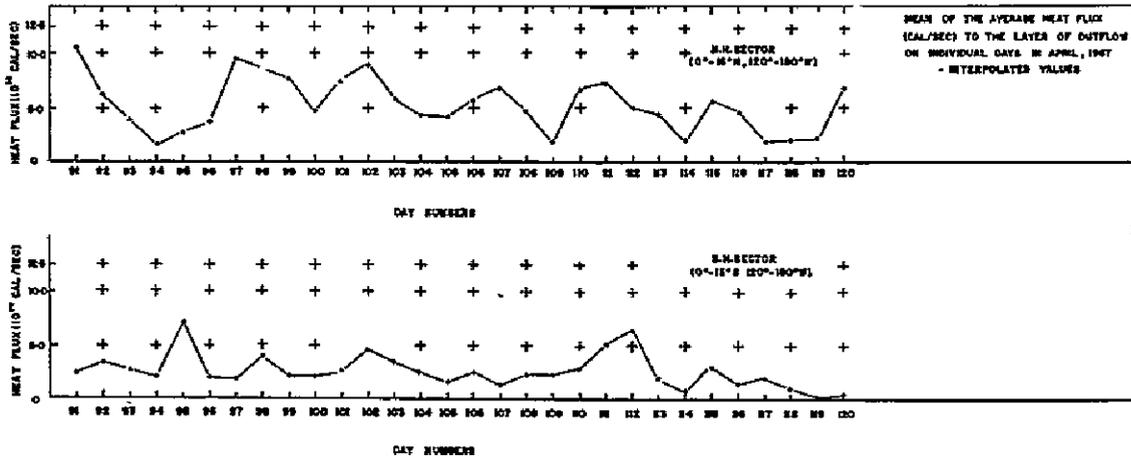


Fig. 35. Daily mean of the average upward heat transport for April, 1967.

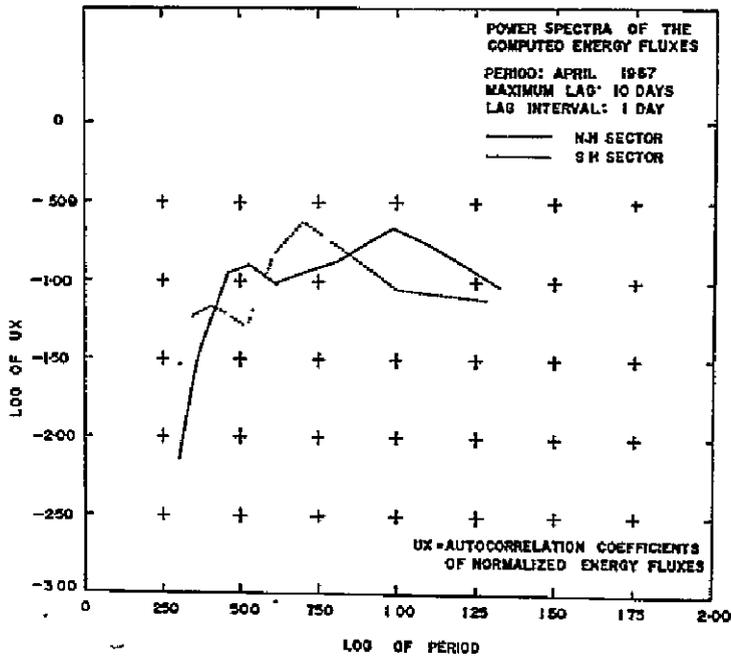


Fig. 36. Power spectra of the computed energy fluxes.

2.8.4 Phase and Speed of the Upper Tropospheric Disturbance Field

In order to study the phase and speed of the upper tropospheric disturbance field, mentioned in the previous section, an analysis was performed in the longitudinally separated grid sections (120°W to 150°W , 150°W to 180°W) in the two hemispheric sectors. No definite conclusions could be derived from this analysis for the following reasons:

- (1) the longitudinal coverage was inadequate for the cloud population to be representative,
- (2) the satellite data was inadequate for sufficient sampling of the cloud populations,
- (3) the large scale cloud was almost always present in the two longitudinal sections across 150°W in the N.H. sector and as a result, the evaluated fluxes could not be assigned properly.

The analysis could, however, be performed with the extended longitudinal coverage around the sub-satellite point and selecting the boundary based on cloud census in the grid box. This will be attempted when adequate data are available.

2.8.5 Implications for Modeling the Tropical Atmosphere

One of the key problems in the models of the tropical atmosphere is to account for the heat release by convection—a phenomenon whose space scale is much smaller than the grid mesh used in the model. The statistics from this study have clearly shown that the release of convection is not random but is controlled (at least in the area under investigation) by large scale motion. Thus there is some hope that one can predict the heat release by sub-mesh scale convections from larger scale motion. Perhaps through use of the tool developed in this research and a determination of the motion field using the cloud displacement, one can learn if it is possible to predict the quantity of convective activity. Finally, to complete the story we must find how the heat thus released affects the large scale motion. In principle, it should be possible to tackle this problem also using our technique. We have been able to estimate the mass flux from the lower to upper troposphere due to convection. The cloud motions in the lower and upper portions of this layer ought to give one useful information on the horizontal momentum in each layer. Thus one would be able to measure the vertical momentum transfer of the visible portion of the atmosphere at least. Using radiosonde data, one might be able to assess transfer in the same "visible" portion of the atmosphere. All this is a task for the future.

3.0 SUMMARY

There is no doubt that the convective clouds in the tropics constitute important elements transforming heat, moisture and momentum in the vertical. The primary purpose of this research has been to show that, based on some reasonable assumptions, one can estimate the convective transport of mass and thus energy fluxes to the tropical upper troposphere from the ATS-I satellite cloud photographs. For extracting this information, a technique has been developed in this research.

From the data presented in the first part of this thesis, the technique is a promising one for the quantitative estimation of upward fluxes of mass and energy associated with a mesoscale system in the tropics. Numerical modelers are interested in the character of mass circulation and the related energy release over wide scales of convection. The growing clouds are the energy sources. This technique will help find the spatial distributions of energy sources from the satellite cloud photographs over the vast oceanic regime. This knowledge would eventually be of importance in the understanding of the interactions between the small scale convection regime and the large scale atmospheric circulations in the tropics.

Although the technique has the excellent capability of evaluating mass and energy upward fluxes in the tropical convection regime, there are some obvious limitations. The technique assumes a cirrus outflow layer of thickness 1.0 km based on a few observations (Anderson, 1960; Ludlum, 1966; Borovikov, 1963). When considered over a long period, this assumption may not be unjustified, but one feels that in the light of its potential for probing tropical energetics, a few aircraft reconnaissance flights over the ITCZ may be of great value by providing this information.

This investigation indicates that the meso- to sub-synoptic scale convections are controlled by the large scale circulations by redistributing the heat sources and sinks. This redistribution is continuously going on in the vertical and over the vast area at the same time. The tropical atmosphere exhibits pronounced short period fluctuations in the convective activity during the day, but not on the diurnal cycle. The upward transport of mass and energy is accomplished through various scales of eddies in the form of high energy surges. On occasion, the amplitudes of these short period pulses may be as high as 1-2 ly/min. These short period pulses have been observed mainly in large cloud ensembles (size larger than 10^{14} cm²).

The spectral estimate of the convective heat fluxes indicate that the convective heat transport over the tropical mid-Pacific pulsates with an approximate periodicity of five days, especially over the N.H. sector. This invariably implies that the convective heat release is controlled by the large

scale motion field. In the S.H. sector, however, a periodicity of convective cloud intensification ranging from 6 to 10 days is indicated. This difference in the periodicity may be attributed to lesser convective cloud population in the S.H. sector in the period of study. Further research in this area will be fruitful.

4.0 PROPOSED FURTHER INVESTIGATION WITH THE ATS-I PHOTOGRAPHS

The magnitude of cloud divergences in the layer of outflow is as well an indication of the character of average vertical motion field in the embedded "hot towers." With the available upward mass flux measurements based on the technique, one can, indeed, estimate the average vertical motion field provided the areas of convergence are known by some direct or indirect means. The radar echoes would indicate the zones of convergence directly beneath the concerned cloud complex but, over the vast oceanic regime in the tropics, such direct measurements are scarce. Therefore, an indirect method needs to be developed utilizing high precision satellite cloud photographs.

It is likely that from the brightness gradients on the digital displays, one will be able to estimate the size of the convergence field in the regime of deep tropical convections. Once this technique is developed, it will be possible to relate the average vertical motion field to the average rainfall rate in the cloud complex for these parameters have unique correlation.

Moreover, the outflow of air at the top of a deep convection originates from the oceanic boundary layer. So, based on some modeling assumptions one might be able to sense the average magnitudes of evaporative fluxes in the oceanic surface layer.

This investigation has indicated that the large scale convective transport to the tropical upper troposphere pulsates with a periodicity of 4 to 5 days. The large scale heat flux from the boundary layer is obviously tied to the wind convergence. The question whether the heat convergence leads or lags the wind convergence in the large scale disturbance field or they appear simultaneously remains to be answered. A geosynchronous meteorological satellite, together with the observations planned in the GARP, may give clues to its answer.

ACKNOWLEDGMENT

I wish to express my sincere gratitude to Professor Verner E. Suomi for his continuous encouragement and patient guidance. Without his assistance and inspiration I would not have succeeded in this research program.

The helpful comments and suggestions of Professors H. H. Lettau, C. E. Anderson and J. Young of the Department of Meteorology were greatly appreciated. I would also like to express my sincere appreciation to Drs. Stephen K. Cox and Thomas H. Vonder Haar for assistance in procuring some of the data used in this dissertation.

But for the understanding and constant encouragement of my wife, Maya, in this difficult venture of mine, I would not have been successful.

I am also thankful to the Government of India for granting me leave of absence from the India Meteorological Department.

This research was supported partly by the U. S. Weather Bureau under grant #27 and by NASA under grant #11542.

REFERENCES

- Ackerman, B., 1959: "The variability of the water contents of tropical cumuli," J. Meteor., 16, 191-198.
- Anderson, C. E., 1960: A study of the pulsating growth of cumulus clouds, Geophysical Research Papers, #72, AFCRL, 136.
- Asai, T. and Kasahara, A., 1967: A theoretical study of the compensating downward motion associated with cumulus clouds, J. Atmos. Sci., 24, 487-496.
- Bartman, F. L., 1967: The reflectance and scattering of solar radiation by the earth, NASA Report, #NAS-54 (03), February, 1967.
- Bates, F. C., 1961: The great plains squall line thunderstorm: A model, Ph.D. dissertation, St. Louis University.
- Borovikov, A. M., et al.: Cloud Physics, U. S. Dept. of Commerce, Washington, 1963.
- Braham, R. R., Jr., 1952: The water and energy budgets of the thunderstorm and their relation to thunderstorm development, J. Meteor., 9, 227-242.
- Brown, R. A., 1966: Three dimensional growth characteristics of an aerographic thunderstorm system. Satellite and Meso-meteorology Research Project Research Paper No. 61, Department of the Geophysical Sciences, University of Chicago, 454.

- Brown, R. A., 1967: Mass and available energy in growing convective clouds, J. Atm. Sci., 24, 308-311.
- Byers, H. and Braham, R., 1949: Thunderstorm, U. S. Dept. of Commerce.
- Conover, J. H., 1962: Cloud Interpretation from Satellite Altitudes, U. S. A. F. Camb. Res. Lab., Res. Note 81, 77, AFCRL-62-680.
- Darkow, G., 1963: A study of infrared radiation measurements in the vicinity of sub-tropical troposphere. Ph.D. thesis, University of Wisconsin.
- Fujita, T., 1963: Analytical Mesometeorology: A Review, Meteor. Monogr., 5, 27, 77-125.
- Green, J. S. A., Ludlum, F. H., and McIlveen, J. F. R., 1966: Isentropic relative flow analysis and the parcel theory, Q. J. Met. Soc., p. 210.
- House, F., 1965: The radiation balance of the earth from a satellite, Ph.D. thesis, University of Wisconsin.
- Lilly, D. K., 1962: On the numerical simulation of buoyant convection, Tellus, 14, 148-172.
- London, J., 1957: Dept. of Meteor., Oceanographic New York Univ., Final Report Project No. 131, Contract No. AF 19(122)-(165).
- Levine, J., 1959: Spherical vortex theory of bubble like motion in cumulus clouds. Woods Hole Oceanographic Inst. Report.
- Ludlum, F. H., 1966: Cumulus and Cumulonimbus Convection, Tellus, 18, 4, 687-698.
- Malkus, J. S. and Ranne, C., 1954: On the structure of some cumulonimbus clouds which penetrated the high tropical troposphere, Tellus, 6, 351-366.
- Malkus, J. S. and Scorer, R. S., 1955: The erosion of cumulus towers, J. Met., 12, 43-57.
- Malkus, J. S., Ronne, C., and Chaffe, M., 1961: Cloud patterns in hurricane Daisy, 1958, Tellus, 13, 8-30.
- Malkus, J. S., 1955: On the formation and structure of downdraft in cumulus clouds, J. Met., 12, 350-354.

- Malkus, J. S., 1960: Penetrative convection and an application to Hurricane Cumulonimbus Towers, Contribution #1044, Woods Hole Oceanographic Institute.
- Malkus, J. S. and Riehl, H., 1964: Cloud structure and distribution over tropical Pacific Ocean, Tellus, 16, 275-287.
- Martin, D., 1968: Report on the 1st session of the study group on tropical disturbances, Joint GARP Organizing Committee.
- Newton, C. W., 1963: Dynamics of severe convective storms, Meteor. Monogr., 5, 27, 33-58.
- Ogura, Y., 1963: The evolution of a moist convective element in a shallow, conditionally unstable atmosphere: A numerical calculation, J. Atmos. Sci., 20, 407-424.
- Orville, 1965: A photogrammetric study of the initiation of cumulus clouds over mountainous terrain, J. Atmos. Sci., 22, 684-699.
- Palmer, C. E., 1951: Tropical Meteorology. Compendium of Meteorology, Amer. Meteor. Soc. 859-880.
- Palmer, C. E., 1952: Tropical Meteorology, Quart. J. Roy. Meteor. Soc., 78, 126-163.
- Riehl, H., et al., 1951: The north-east trade of the Pacific Ocean, Quart. J. Roy. Meteor. Soc., 77, 598-626.
- Riehl, H. and Malkus, J. S., 1957: On the heat balance and maintenance of circulation in the trades, Quart. J. Roy. Meteor. Soc., 83, 21-29.
- Riehl, H. and Malkus, J. S., 1961: Some aspects of Hurricane Daisy, Tellus, 13, 181-213.
- Riehl, H. and Malkus, J. S., 1958: On the heat balance in the equatorial trough zone, Geophysics, 6, 503-538.
- Sikdar, D. N., 1968: Report on the 1st session of the study group on tropical disturbances. Joint GARP Organizing Committee.
- Suomi, V. E. and Fujita, T. T., 1968: Report on the 1st session of the study group on tropical disturbances. Joint GARP Organizing Committee.
- Stommel, H., 1947: Entrainment of air into cumulus cloud, J. Met., 8, 127-129.

- Vonnegut, B., C. B. Moore and Botka, A. T., 1959: Preliminary results of an experiment to determine initial precedence of organized electrification and precipitation in thunderstorms, J. Geophys. Res., 64, 347-357.
- Vonder Haar, Thomas, 1968: A study of the variations of the global radiation budget from an extended time series of satellite radiation data. Ph.D. thesis, University of Wisconsin.
- Whitney, L. F., 1963: Severe storm clouds as seen from TIROS, J. Appl. Meteor., 2, 501-507.
- Yanai, M., et al., 1968: Power spectra of large-scale disturbances over the tropical Pacific, J. Met. Soc., Japan, 46, 801-816.
- Zipser, E. J. and Taylor, B. C., 1968: A catalog of meteorological data obtained during the Line Island Experiment, Feb.-April, 1967. NCAR TN-35, National Center for Atmospheric Research, Boulder, 362 pp.

APPENDIX I

TABLE I
 Measurements on ATS-I Satellite Photographs
 Day 109

Cloud No.	Area I (10^{12}cm^2)	Area 2 (10^{12}cm^2)	$A_2 - A_1$ (10^{12}cm^2)	Mean area (10^{12}cm^2)	Cloud Div. (10^{-4} per sec)
1	81.280	99.500	18.220	90.390	1.623
2	7.516	6.955	-0.561	7.236	-0.540
3	1.049	1.835	0.787	1.442	5.436
4	1.486	1.932	0.446	1.709	2.174
5	2.272	2.318	0.046	2.295	0.146
6	7.604	7.728	0.124	7.666	0.118
7	5.069	7.921	2.852	6.495	4.076
8	3.496	3.671	0.175	3.583	0.362
9	3.234	3.478	0.244	3.356	0.543
10	26.220	27.050	0.828	26.630	0.228
11	51.390	71.480	20.090	61.440	2.832
12-13	68.870	73.130	4.255	71.000	0.447
14	11.890	14.970	3.087	13.430	1.881
15	4.370	2.512	-1.858	3.441	-3.080
16	3.496	1.932	-1.564	2.714	-3.241
17	8.740	5.796	-2.944	7.268	-2.440
18	3.321	2.705	-0.616	3.013	1.343
19	9.614	11.590	1.978	10.600	1.490
20	637.8	725.10	87.30	681.5	0.991

PRECEDING PAGE BLANK NOT FILMED

TABLE II
Measurements on ATS-I Satellite Photographs
Day 109

Cloud No.	Area 1 (10^{12}cm^2)	Area 2 (10^{12}cm^2)	$A_2 - A_1$ (10^{12}cm^2)	Mean area (10^{12}cm^2)	Cloud Div. (10^{-4} per sec)
1	99.500	107.300	7.753	103.400	0.564
2	6.955	11.490	4.536	9.223	4.725
3-4	3.767	3.830	0.063	3.799	0.121
5	2.318	2.298	-0.020	2.308	-0.062
6	7.728	9.576	1.848	8.652	1.732
7	7.921	9.097	1.176	8.509	1.075
8	3.671	3.830	0.160	3.751	0.315
9	3.478	3.926	0.449	3.702	0.934
10	27.050	32.560	5.510	29.800	1.476
11	71.480	84.270	12.780	77.880	1.294
12-13	73.130	98.630	25.510	85.880	2.527
14	14.970	12.450	-2.524	13.710	-1.221
15	2.512	4.213	1.702	3.363	4.909
16	1.932	4.788	2.856	3.360	10.710
17	5.796	6.129	0.333	5.962	0.416
18	2.705	6.320	3.615	4.512	9.680
19	11.590	11.490	-0.101	11.540	-0.063
20	725.10	760.60	35.50	742.9	0.354

TABLE III

Cloud No.	Volume flux (V_t) at the cloud top (10^{14} cc/sec)	Mass flux (M_D) from the subcloud layer (10^8 kg/sec)	Heat flux (E_C) (10^{11} cal/sec)
1	13.20	3.57	38.40
2			
3	0.57	0.15	1.66
4	0.32	0.09	0.94
5	0.03	0.01	0.10
6	0.09	0.02	0.26
7	2.06	0.56	6.01
8	0.13	0.03	0.37
9	0.18	0.05	0.51
10	0.60	0.16	1.75
11	14.50	3.94	42.40
12-13	3.07	0.84	8.97
14	2.23	0.61	6.51
15			
16			
17			
18			
19	1.43	0.39	4.17
20	56.75	15.42	165.70

TABLE IV

Cloud No.	Volume flux(V_t) at the cloud top (10^{14} cc/sec)	Mass flux (M_D) from the subcloud layer (10^8 kg/sec)	Heat flux (E_C) (10^{11} cal/sec)
1	5.56	1.51	16.20
2	3.25	0.88	9.50
3-4	0.05	0.01	0.13
5			
6	1.32	0.36	3.87
7	0.84	0.23	2.46
8	0.11	0.03	0.33
9	0.32	0.09	0.94
10	3.95	1.07	11.50
11	9.16	2.49	26.80
12-13	18.30	4.97	53.40
14			
15	1.22	0.33	3.56
16	2.05	0.56	5.98
17	0.24	0.06	0.70
18	2.59	0.70	7.57
19			
20	23.00	6.25	7.15

TABLE V

Using Brown's (1967) Energy Equation

$$M_A = \left(\frac{E_C}{1.4 \times 10^5} \right)^{1/1.4} \text{ (kg.)}$$

and our Heat Flux Data, E_C

Cloud No.	From Table III		From Table IV	
	Energy (E_C) (10^{21} ergs)	Cloud Mass (M_A) (10^{11} kg)	Energy (E_C) (10^{21} ergs)	Cloud Mass (M_A) (10^{11} kg)
1	223.00	28.70	94.80	13.60
2			55.50	8.47
3	9.62	1.82		
4	5.45	1.11	0.77	0.20
5	0.56	0.15		
6	1.52	0.36	22.60	3.85
7	34.90	5.64	14.40	2.59
8	2.14	0.49	1.95	0.45
9	2.98	0.65	5.48	1.11
10	10.10	1.91	67.40	10.00
11	246.00	31.30	156.00	21.00
12-13	52.00	8.01	312.00	38.50
14	37.70	6.04		
15			20.80	3.58
16			34.90	5.64
17			4.07	0.86
18			44.20	6.94
19	24.20	4.09		
20	960.50	94.80	392.10	432.60

APPENDIX 2

PRECEDING PAGE BLANK NOT FILMED

TABLE I

Area of Clouds for Day 107
(10^{14} cm²)

NORTHERN HEMISPHERE

(Longitude 120°W - 180°W, Latitude 0° - 15°N)

Cloud No.	Picture 1 <u>194802Z</u>	Picture 2 <u>201115Z</u>	Picture 3 <u>203430Z</u>	Picture 4 <u>205805Z</u>	Picture 5 <u>212101Z</u>	Picture 6 <u>214415Z</u>
1	2.8	4.4	1.3	3.4	3.7	3.8
2	1.5	2.2	0.2	1.1	1.4	2.4
3	26.0	33.0	24.0	2.6	24.0	25.0
4	7.2	75.0	65.0	72.0	69.0	69.0
5	0.6	5.7	2.6	2.7	0.9	1.6
6	0.2	0.3	0.1	0.6	1.0	1.3
7	200.0	230.0	160.0	190.0	150.0	150.0
8	4.8	5.9	3.4	1.7	4.7	5.3
Totals	240.0	350.0	260.0	280.0	250.0	260.0
% of cloudiness	21.9%	31.9%	23.1%	25.0%	22.9%	23.3%

Cloud No.	Picture 7 <u>220735Z</u>	Picture 8 <u>223046Z</u>	Picture 9 <u>225400Z</u>	Picture 10 <u>231715Z</u>	Picture 11 <u>234029Z</u>	Picture 12 <u>240349Z</u>
1	4.1	4.6	4.1	3.7	5.6	5.6
2	3.0	4.1	2.9	2.7	2.4	2.6
3	22.0	25.0	19.0	23.0	27.0	26.0
4	70.0	78.0	78.0	85.0	110.0	120.0
5	1.6	1.9	2.0	2.8	4.7	4.7
6	1.4	1.6	1.5	2.3	4.3	4.1
7	160.0	150.0	130.0	150.0	76.0	180.0
8	5.5	6.8	6.2	6.6	6.0	6.2
Totals	270.0	280.0	240.0	280.0	230.0	340.0
% of cloudiness	24.1%	24.9%	22.0%	25.4%	20.8%	31.1%

Mean of percent - 24.7%

PRECEDING PAGE BLANK NOT FILMED

Table I (continued)

<u>SOUTHERN HEMISPHERE</u>						
(Longitude 120°W - 180°W, Latitude 0° - 15°S)						
<u>Cloud No.</u>	<u>Picture 1 194802Z</u>	<u>Picture 2 201115Z</u>	<u>Picture 3 203430Z</u>	<u>Picture 4 205805Z</u>	<u>Picture 5 212101Z</u>	<u>Picture 6 214415Z</u>
1	9.4	12.0	12.0	12.0	12.0	14.0
2	13.0	15.0	13.0	15.0	10.0	11.0
3	3.7	4.0	2.7	3.1	2.2	1.4
4	2.5	2.3	3.3	1.6	2.4	2.1
5	2.3	1.4	2.1	2.7	2.3	1.7
6	2.3	2.5	1.0	1.2	1.3	1.1
7	3.6	2.7	2.1	1.8	2.4	2.4
8	19.0	26.0	14.0	23.0	22.0	22.0
Totals	56.0	67.0	50.0	60.0	55.0	56.0
% of cloudiness	5.1%	6.0%	4.5%	5.5%	5.0%	5.0%
<u>Cloud No.</u>	<u>Picture 7 220735Z</u>	<u>Picture 8 223046Z</u>	<u>Picture 9 225400Z</u>	<u>Picture 10 231715Z</u>	<u>Picture 11 234029Z</u>	<u>Picture 12 240349Z</u>
1	14.0	17.0	16.0	17.0	23.0	23.0
2	8.7	10.0	7.7	9.0	13.0	15.0
3	1.3	1.6	1.7	1.8	1.7	1.9
4	1.7	2.3	2.1	1.4	3.8	6.6
5	2.5	2.6	3.5	3.5	6.1	6.9
6	1.3	1.2	1.3	1.1	2.2	2.9
7	1.6	2.6	2.7	1.8	0.8	0.7
8	23.0	24.0	23.0	26.0	24.0	30.0
Totals	55.0	62.0	58.0	62.0	74.0	87.0
% of cloudiness	4.9%	5.6%	5.3%	5.6%	6.7%	7.9%
Mean of percent - 5.6%						

TABLE II
 Cloud Divergence for Successive Time Intervals, Day 107
 (10^{-4} per sec)

NORTHERN HEMISPHERE

(Longitude $120^{\circ}\text{W} - 180^{\circ}\text{W}$, Latitude $0^{\circ} - 15^{\circ}\text{N}$)

<u>Cloud No.</u>	<u>1st Period</u>	<u>2nd Period</u>	<u>3rd Period</u>	<u>4th Period</u>	<u>5th Period</u>	<u>6th Period</u>
1	4.1	-5.1	12.0	0.7	0.1	0.6
2	3.6	-6.4	25.0	2.4	4.7	2.1
3	2.2	-2.0	-6.3	61.0	0.1	-0.8
4	68.0	-1.0	0.8	-0.3	0.0	0.1
5	60.0	-3.9	0.4	-5.0	6.3	0.0
6	3.6	-3.6	23.0	5.3	1.9	0.6
7	0.9	-2.1	1.5	-1.7	0.0	0.4
8	1.7	-3.1	-3.6	13.0	0.8	0.3

<u>Cloud No.</u>	<u>7th Period</u>	<u>8th Period</u>	<u>9th Period</u>	<u>10th Period</u>	<u>11th Period</u>
1	0.9	-0.7	-7.6	3.7	-1.2
2	2.6	-2.2	-0.3	-0.8	0.5
3	0.7	-1.5	1.4	1.2	-0.3
4	0.7	0.0	0.7	1.6	0.9
5	1.4	0.4	2.9	4.6	0.0
6	1.0	-0.4	3.4	6.4	0.3
7	-0.2	-1.2	1.4	-3.7	9.7
8	1.7	-0.6	0.5	-7.0	0.2

Table II (continued)

<u>SOUTHERN HEMISPHERE</u>						
(Longitude 120°W - 180°W, Latitude 0° - 15°S)						
<u>Cloud No.</u>	<u>1st Period</u>	<u>2nd Period</u>	<u>3rd Period</u>	<u>4th Period</u>	<u>5th Period</u>	<u>6th Period</u>
1	2.1	0.0	-1.8	-0.1	1.2	0.3
2	1.4	-1.2	1.2	-2.2	0.3	-1.4
3	0.5	-2.3	0.9	-2.1	-2.5	-0.8
4	-0.5	2.9	-3.7	3.7	-0.9	-1.3
5	-2.9	3.3	2.1	-1.1	-1.8	3.4
6	0.6	-4.4	1.6	0.7	-1.1	1.3
7	-1.8	-1.6	-0.7	2.0	0.0	-2.3
8	2.6	-3.4	4.7	-0.2	0.0	0.3
<u>Cloud No.</u>	<u>7th Period</u>	<u>8th Period</u>	<u>9th Period</u>	<u>10th Period</u>	<u>11th Period</u>	
1	1.5	-0.5	0.5	2.4	0.1	
2	1.3	-1.8	1.2	3.2	1.2	
3	2.2	-0.4	0.3	-0.5	0.8	
4	2.6	-0.6	-2.3	12.0	5.3	
5	0.3	2.4	0.0	5.3	1.0	
6	-1.0	1.2	-1.1	6.4	2.5	
7	4.4	0.3	-2.4	-3.8	-0.8	
8	0.1	-0.1	0.9	-0.7	1.9	

TABLE III

Mass Flux from the Layer of Inflow, Day 107
(10^{12} g/sec)

NORTHERN HEMISPHERE(Longitude $120^{\circ}\text{W} - 180^{\circ}\text{W}$, Latitude $0^{\circ} - 15^{\circ}\text{N}$)

<u>Cloud No.</u>	<u>1st Period</u>	<u>2nd Period</u>	<u>3rd Period</u>	<u>4th Period</u>	<u>5th Period</u>	<u>6th Period</u>
1	2.8		3.7	0.6	0.1	0.6
2	1.3		1.4	0.6	1.6	1.2
3	14.0			39.0	0.6	
4	120.0		14.0		0.0	1.6
5	9.0		0.3		1.3	0.0
6	0.2		0.8	0.8	0.5	0.2
7	45.0		57.0		2.2	16.0
8	2.0			5.5	1.0	0.5
Totals	190.0		77.0	47.0	7.4	20.0

<u>Cloud No.</u>	<u>7th Period</u>	<u>8th Period</u>	<u>9th Period</u>	<u>10th Period</u>	<u>11th Period</u>
1	0.9			3.4	
2	1.9				0.3
3	4.1		6.7	7.0	
4	13.0	1.4	13.0	35.0	22.0
5	0.5	0.2	1.5	3.2	
6	0.4		1.3	3.6	
7			45.0		180.0
8	2.3		0.7		0.3
Totals	23.0	1.5	68.0	52.0	200.0

Mean of
the totals - 63.0

Table III (continued)

SOUTHERN HEMISPHERE
(Longitude 120°W - 180°W, Latitude 0° - 15°S)

<u>Cloud No.</u>	<u>1st Period</u>	<u>2nd Period</u>	<u>3rd Period</u>	<u>4th Period</u>	<u>5th Period</u>	<u>6th Period</u>
1	4.9	0.1			3.4	1.0
2	4.4		3.7		0.8	
3	0.5		0.6			
4		1.6		1.4		
5		1.2	1.1			1.4
6	0.3		0.4	0.2		0.4
7				1.0		
8	13.0		16.0			1.9
Totals	23.0	2.9	22.0	2.6	4.1	4.6

<u>Cloud No.</u>	<u>7th Period</u>	<u>8th Period</u>	<u>9th Period</u>	<u>10th Period</u>	<u>11th Period</u>
1	5.3		2.1	10.0	0.8
2	2.7		2.3	7.1	3.8
3	0.7	0.2	0.1		0.3
4	1.1			4.2	4.9
5	0.2	1.5	0.1	4.5	1.5
6		0.3		1.8	1.3
7	1.7	0.2			
8	0.7		4.9		11.0
Totals	12.0	2.2	9.6	28.0	24.0

Mean of
the totals - 12.0

Table IV (continued)

<u>SOUTHERN HEMISPHERE</u>						
(Longitude 120°W - 180°W, Latitude 0° - 15°S)						
<u>Cloud No.</u>	<u>1st Period</u>	<u>2nd Period</u>	<u>3rd Period</u>	<u>4th Period</u>	<u>5th Period</u>	<u>6th Period</u>
1	5.2	0.1			3.6	1.0
2	4.7		4.0		0.8	
3	0.5		0.6			
4		1.8	1.5			
5		1.2	1.2			1.5
6	0.3		0.4	0.2		0.4
7				1.0		
8	13.0		17.0			2.0
Totals	24.0	3.1	23.0	2.8	4.4	5.0
Heat flux per sq. cm.	2.2	0.3	2.1	0.3	0.4	0.5
<u>Cloud No.</u>	<u>7th Period</u>	<u>8th Period</u>	<u>9th Period</u>	<u>10th Period</u>	<u>11th Period</u>	
1	5.7		2.3	11.0	0.8	
2	2.9		2.5	7.7	4.1	
3	0.7	0.2	0.1		0.3	
4	1.2			4.5	5.3	
5	0.2	1.6	0.1	4.9	1.6	
6		0.4		1.9	1.4	
7	1.9	0.2				
8	0.8		5.3		12.0	
Totals	13.0	2.4	10.0	30.0	25.0	
Mean of the total	15.2					
Heat flux per sq. cm.	1.2	0.2	0.9	2.7	2.3	
Mean of the total heat flux per sq. cm. (10^{-3})	- 1.2					

N71-11617

A CENSUS OF CLOUD SYSTEMS IN THE TROPICAL PACIFIC OCEAN¹

by

Otto J. Karst

ABSTRACT

As a supporting study to a GARP tropical meteorology experiment a census was made on cloud systems in the tropical Pacific Ocean using ATS I and ESSA III satellite photographs. The study area ranged from 110°E to 100°W and 25°S to 25°N, and was subdivided into five equal regions. Organized cloud systems were classified by geometric patterns into an oval, line, wave, or spiral vortex. There are large differences between the northern and southern hemisphere in the level of organization and the frequency of cloud systems. The northcentral Pacific region (150°W to 160°E) has a lower level of organization than the two other regions in the northern hemisphere. The study shows that there are high frequencies of waves and vortices for the northeast (150°W to 100°W) and northwest (160°E to 110°E) tropical Pacific regions during the summer and early fall months. In terms of the frequency of organized cloud systems the northeast region and the northwest region would seem to be about equally suitable for a prospective tropical field experiment.

¹Note added in proof: Since the completion of this study, an extension to the tropical Atlantic has been supported by the same contract; in addition, inter-comparison of the present paper and the work by Vonder Haar *et al.*, included in this same report, has purposely been avoided up to the present time.

I. Introduction

The impetus for a cloud census of the tropical Pacific Ocean arose from the GARP joint organizing committee in Stockholm, 1967. At this time the committee was planning a detailed tropical meteorology experiment in the Pacific Ocean. The purpose of the study was to investigate the interaction of the cloud cluster scale phenomenon with that of the wave scale. Recommendations were made for a preliminary census of cloud systems in this area so that the best location for the study could be selected. The location that is chosen would have to contain a large number of organized cloud systems throughout the study period.

II. Data Used

Martin and Karst (1968) conducted an initial cloud census of the central and eastern Pacific Ocean from 160°E to 100°W, and 25°S to 25°N. The period of study was from March, 1967 to February, 1968, inclusive. The satellite pictures available for the census were ATS I geosynchronous and ESSA III photographs.

On the suggestion of the GARP tropical working group a supplementary cloud census in the northwest Pacific Ocean from 160°E to 110°E, and 0° to 25°N was added to the basic census. Only ESSA III photographs were available for this region.

Figure 1 shows how the study areas was subdivided into five equal regions: northeast, northcentral, northwest, southcentral, and southeast.

III. Method of Investigation

Each day's satellite photographs were examined for organized cloud systems. These systems were tracked from day to day until they dissipated, merged with other systems, or moved out of the major boundary of the study area. The systems were classified according to a geometric appearance: oval, line, wave, and spiral vortex. Figures 2, 3, 4, and 5 give representative examples of each cloud category, respectively.

"Oval" was comparable to a cloud cluster because it contained a mass of cloud with no observable rotation. A typical cumulonimbus would be classified as an oval. "Line" classification indicated a cloud system stretched out along an axis. A subtropic frontal system that extended into the tropics would be classified as a line. The ITC in itself was not classified as a line, but disturbances on the ITC were included in the statistics. At times it was difficult to separate a line and oval classification since the geometric shape of the cloud system was the only guide used. "Wave" represented any disturbance

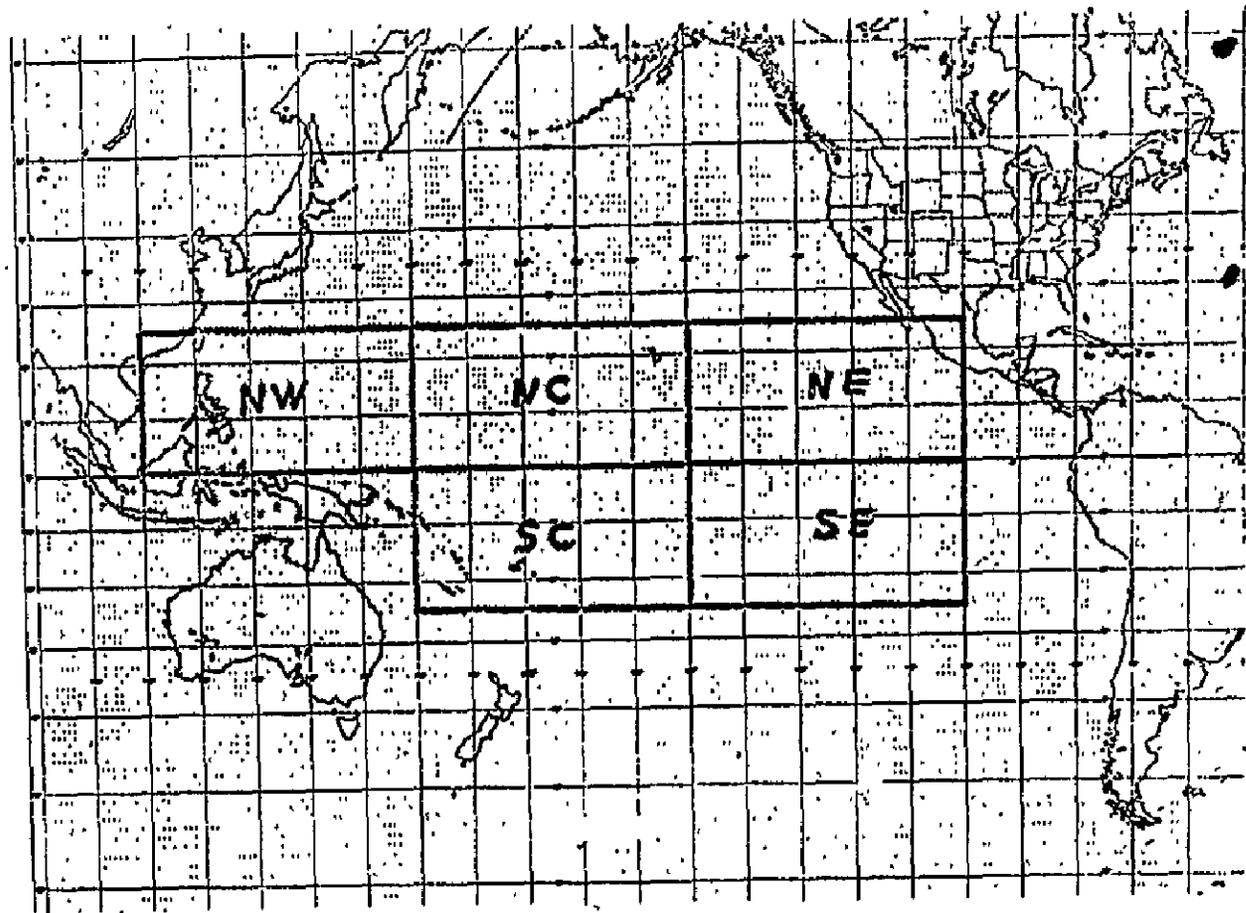


Fig. 1. Five regions in the tropical Pacific Ocean included in the cloud census.

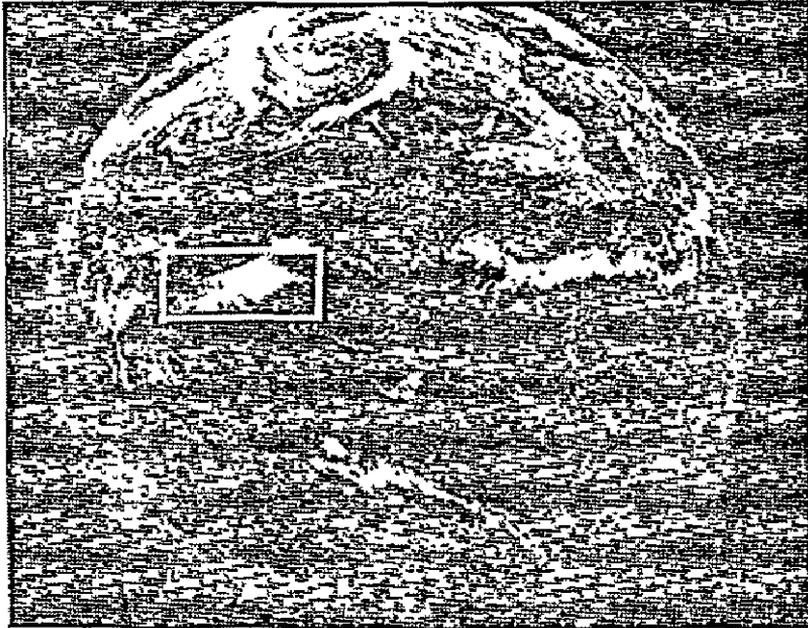


Fig. 2. ATS I photograph with a geometric oval pattern outlined in white.

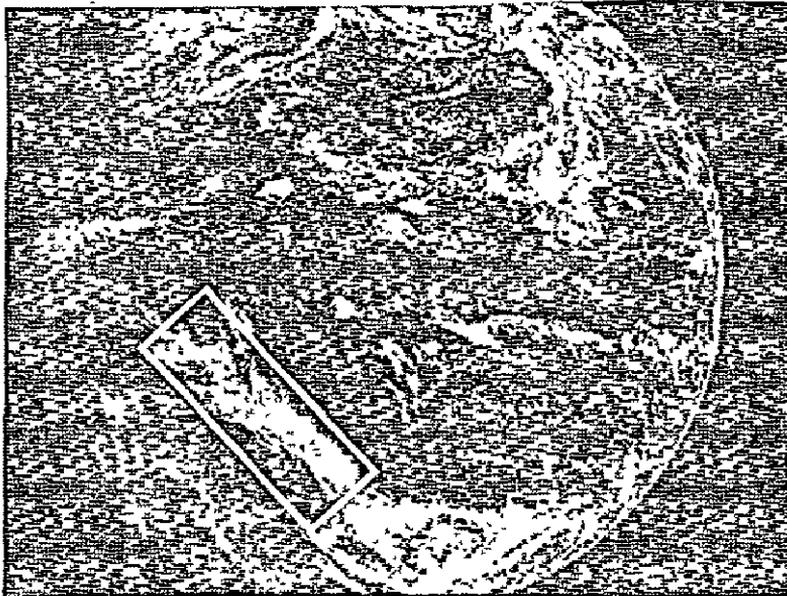


Fig. 3. ATS I photograph with a geometric line pattern outlined in white.

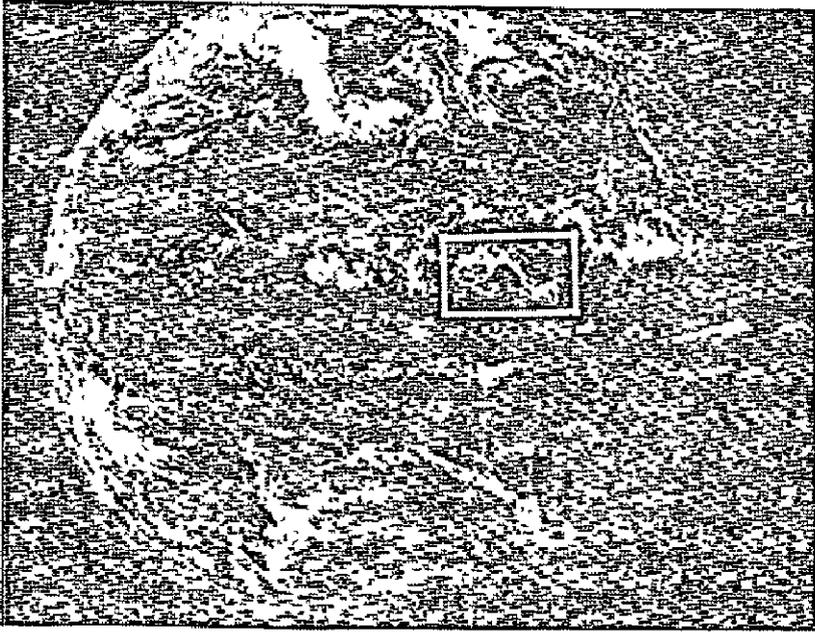


Fig. 4. ATS I photograph with a geometric wave pattern outlined in white.

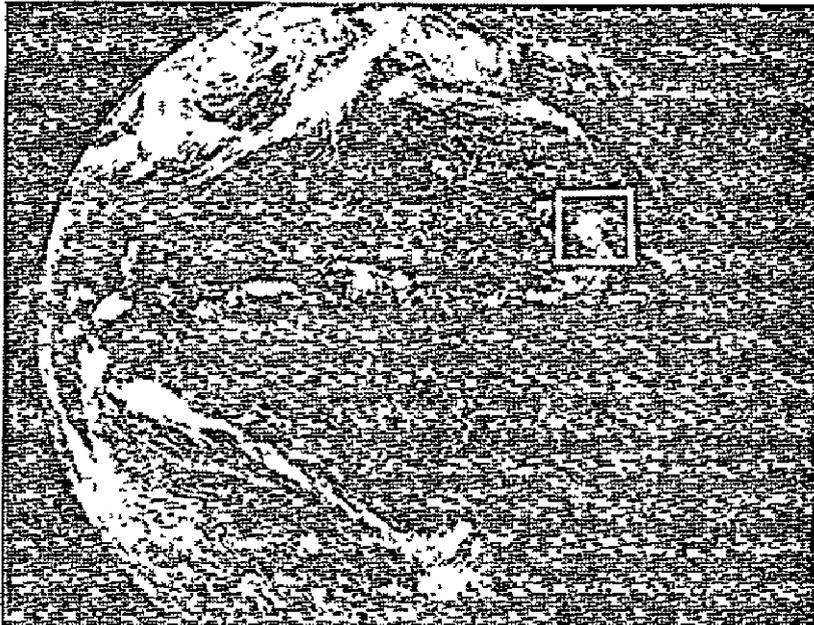


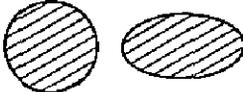
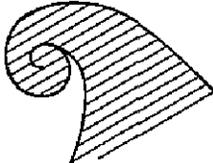
Fig. 5. ATS I photograph with a geometric spiral vortex pattern outlined in white.

with an inverted V geometric design. Typically this system was observed along the ITC. "Vortex" was used for any system that had rotation.

The motivation behind this geometric classification was developed through personal correspondence with E. G. Astling. See Fig. 6 for the development of a spiral from a line.

In addition to the geometric classification, the following parameters were also noted: 1) degree of organization (weak, moderate, pronounced), 2) location (latitude and region), 3) size (cloud area in degrees of latitude squared), 4) growth from the previous day, 5) movement of the system in 24 hours, 6) and lifetime of the system. Also the parameters of cloud height and direction of blowoff were estimated from the satellite photographs.

FORMS OF TROPICAL CLOUD DISTURBANCES

TERM	IDEALIZED SHAPE (NORTHERN HEMISPHERE)	LINEARIZED FORM
OVAL		
LINE		
WAVE		
SPIRAL VORTEX		



(Originated by E G Astling)

Fig. 6. Development of a spiral from a line.

ATS time lapse films were useful in establishing the original classification of the individual parameters mentioned above. The films were particularly good in estimating degree of organization, growth, movement, and blowoff. There was good agreement in data when comparing the original estimate of cloud parameters derived from individual ATS I cloud photographs with those ATS I photographs available in film series for April 1967.

IV. Results of the Cloud Census

The data was analyzed to study monthly, seasonal, and regional variations in the observed level of organization of cloud systems. Graphs were drawn to show the differences in the frequency of occurrence for each level of organization of cloud systems. In addition graphs were prepared to show the variations in the other recorded cloud parameters: mean movement, length of life, and cloud area.

I. Daily Frequency Distribution

Figure 7 represents the frequency of occurrences for the four levels of cloud organization in each month or region. Each system was tallied once each day of its lifetime or until it moved out of the major boundary of the study area. What follows is a description of the major characteristics in each of the five regions.

Northeast Region

A large peak in oval frequency is found in August, and the peak vortex distribution occurs in June, July, August, September, and October. There is a sharp decline in the frequency of vortices in November. Waves reach a maximum in the early fall and winter months of September through January, and lines are also more active in the late fall and winter months.

Northcentral Region

The large number of ovals in the summer months of June, July, and August decrease to a minimum in September, and rise to another maximum in October. The vortex peak in August and September is not nearly as pronounced as in the northeast region. Waves are more frequent in the fall months of September and November, but lines show an increase in frequency in early winter.

The wave maximum which occurs in early fall and winter months in the northeast and northcentral region coincides with the end of the rainy season in the northern hemisphere, and is related to a decline in the vortex distribution and an increase in frequency of lines. There are considerably more ovals in the northcentral region than in the northeast, but the vortex frequency is reduced. During this period the level of organization is lower in the northcentral.

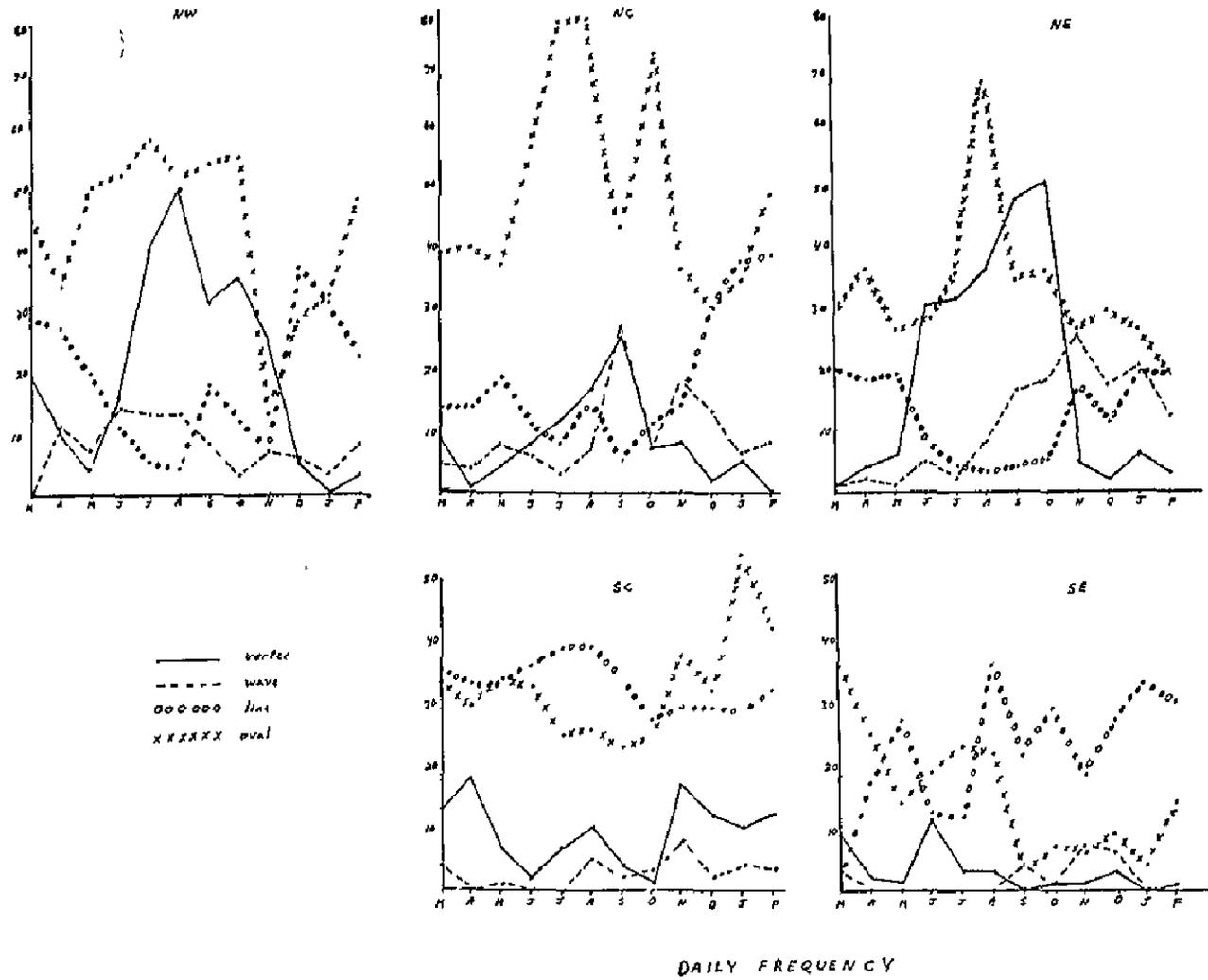


Fig. 7. Daily frequency distribution of cloud systems for each month and region.

Northwest Region

Ovals are fairly active except for a decline in frequency during November. The large peak of vortices in July, August, September, and October is similar to the curves shown for the northeast region. However, the sharp decline in frequency occurs in December. While waves are more frequent in the summer months, lines again are more active in the winter months.

Southcentral Region

An inspection of frequency shows that lines are the most frequent systems in the southern hemisphere winter, and ovals have a sharp increase in activity during the summer months. Vortices fluctuate with a general decline in activity during the winter season. Waves are far less active than in the northern hemisphere. Notice that vortices do not have the large southern hemisphere summer maximum that were observed in the northern hemisphere summer. However, ovals do increase sharply in the summer months. This suggests that the level of organization is lower in the southern hemisphere summer.

Southeast Region

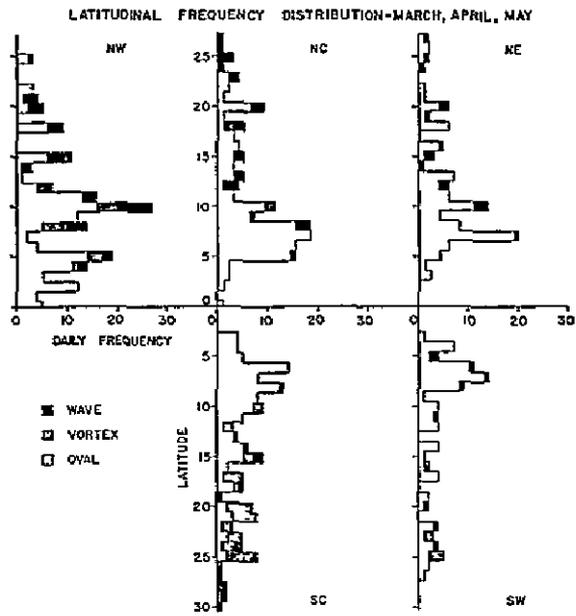
Lines are by far the most active system and ovals show a large decline in frequency during the southern hemisphere summer. This is the most inactive region for waves and vortices.

Large differences in convective activity exist between the northern and southern hemisphere. The cloud averaging techniques of the ESSA III photographs developed by Kornfield and Hasler (1968) in the University of Wisconsin point out clearly the fact that the northern hemisphere is much more active than the southern hemisphere. One sees the persistent line feature in the southcentral region of the tropics and the relatively low level of organization in the cloud systems for the northcentral region of the Pacific Ocean.

Gray (1968) studied regions where typhoons form relying partly on data from pre-satellite days. He showed a large tendency for typhoons to form in the northeast and northwest regions, with virtually no typhoon formation in the northcentral region. The cloud census results agree with Gray's observed vortex distribution in the northern hemisphere.

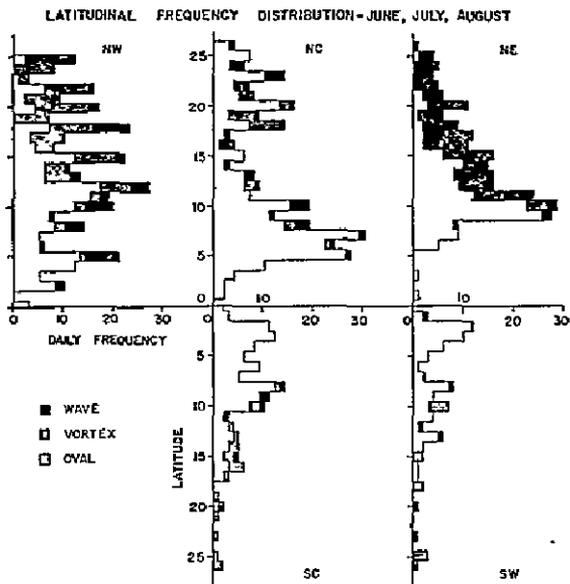
2. Latitude Frequency Distribution

A mean latitude position was assigned to every cloud system for each day it was observed in the study area. Figures 8a, b, c, d represent the seasonal and regional variations in the frequency of occurrence for each level of cloud organization as a function of mean latitude position. As in Figure 7 each system was tallied once each day of its lifetime, but only ovals, waves, and vortices



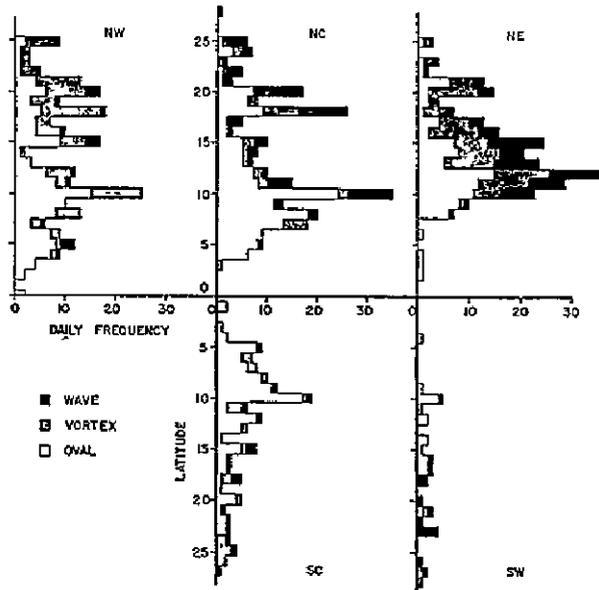
(a) March-April-May

Figs. 8a, b, c, d. Latitude frequency distribution of daily cloud systems for each month and region.



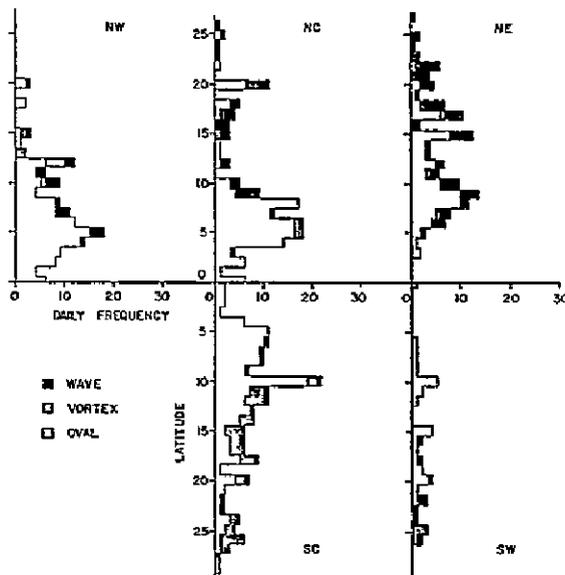
(b) June-July-August

LATITUDINAL FREQUENCY DISTRIBUTION-SEPTEMBER, OCTOBER, NOVEMBER



(c) September-October-November

LATITUDINAL FREQUENCY DISTRIBUTION-DECEMBER, JANUARY, FEBRUARY



(d) December-January-February

are included in the statistics. The seasons, beginning with March, April, and May, are chosen to correspond to the spring, summer, fall, and winter months. What follows is a description of the major characteristics in each of the seasons.

March - April - May

Ovals are distributed evenly in all regions, but notice the distribution from 0° to 5° N in the northwest region. The peak distribution for ovals is observed around $5 - 10^{\circ}$ latitude, except in the northwest region where systems are found close to the equator, but then peak at 10° . There is a peak vortex distribution in the southcentral at $20 - 25^{\circ}$.

June - July - August

The peak distribution for vortices in the northeast and northwest region is found at $15 - 25^{\circ}$ latitude, and equatorward for ovals in the southern hemisphere at $0 - 5^{\circ}$. In the northcentral region notice that the peak distribution for ovals is still between $5 - 10^{\circ}$ latitude.

September - October - November

The peak distribution for vortices is around $10 - 15^{\circ}$ in the northeast, and $15 - 20^{\circ}$ in the northwest. The wave peak in the northeast corresponds to the vortex distribution. There is a noticeable peak in the southcentral region for ovals at $5 - 10^{\circ}$ and no peak occurs in the southeast region where activity is diminished. Ovals peak in the northcentral region at 10° latitude.

December - January - February

Oval activity is found near the equator in the northwest region with a peak distribution at 5° . The peak activity in the northeast and northcentral regions has moved equatorward to $5 - 10^{\circ}$, with peak activity in the southern hemisphere moving poleward to 10° .

In summary one notices that large oval activity is restricted equatorward of 15° latitude in all regions. The peak activity in the northern hemisphere is usually found between $5 - 10^{\circ}$ latitude in the winter and $10 - 15^{\circ}$ in the summer months, except in the northcentral where it is still $5 - 10^{\circ}$. The vortex distribution increases sharply in the northern hemisphere in June - July - August with the systems found at $15 - 25^{\circ}$ latitude. This distribution moves equatorward in the fall months to $10 - 20^{\circ}$ latitude. So there is a shift poleward of the vortex distribution in the summer months, and a shift equatorward in the fall months.

These shifts in the mean latitude distribution can be correlated to the sudden seasonal shifts in the mean position of the ITC. Oval activity is generally found associated with or equatorward of the ITC, while vortices are found poleward of the ITC. The seasonal shifts in the ITC were verified by Kornfield and Hasler (1968) in the photographic cloud averages of the ESSA III satellite photographs.

3. Percent Frequency Distribution

Figure 9 represents the percent of cloud systems in a given region and season that reach each level of organization. A system is tallied once for each level of organization that it attains during its lifetime within the study area. For an example if a cloud progresses from a wave to a vortex, and dissipates as an oval, then it would be classified once as a wave, once as a vortex, and once as an oval.

A vortex maximum is located in the northeast (40% of systems reach this level of organization) and northwest (35%) during the summer months. However, vortex frequency is noticeably smaller but more uniform in the northcentral (10%). Wave maximum is reached in the northeast during the late summer, fall, and winter months (up to 40% of the systems reach this level of organization). The wave peak in the northcentral is only observed in September (25%), while the northwest maximum is in June, July, and August (30%). Lines tend to be more frequent in the winter months, but this is not as noticeable for the northeast region.

4. Mean Movement

Figures 10a, b, c, d represent mean movement of cloud systems in three degree latitudinal strips for each season and region. The vectors point in the direction the system was moving with the scale 1 inch = 10 degrees per day. Movements for a strip were averaged individually for each compass point, i. e.,

the mean movement $\vec{V}_0 = \frac{1}{n_0} \sum_{i=1}^{\infty} \vec{V}_i$ where n_0 is the number of movements in a given direction and V_i is the individual movements in a given direction.

In the northern hemisphere in all three regions we had strong westward movement in June, July, August, September, October, and November. There is a turning to eastward movement in the winter months in the northern hemisphere at 10 - 25°N latitude. In the southern hemisphere in the southcentral region there is the eastward turning at 13 - 25° latitude in all seasons due to the frontal lines that occur in this region.

One might want to compare the movement of cloud systems with observed streamline analysis in the central and western Pacific Ocean at various pressure levels in the atmosphere.

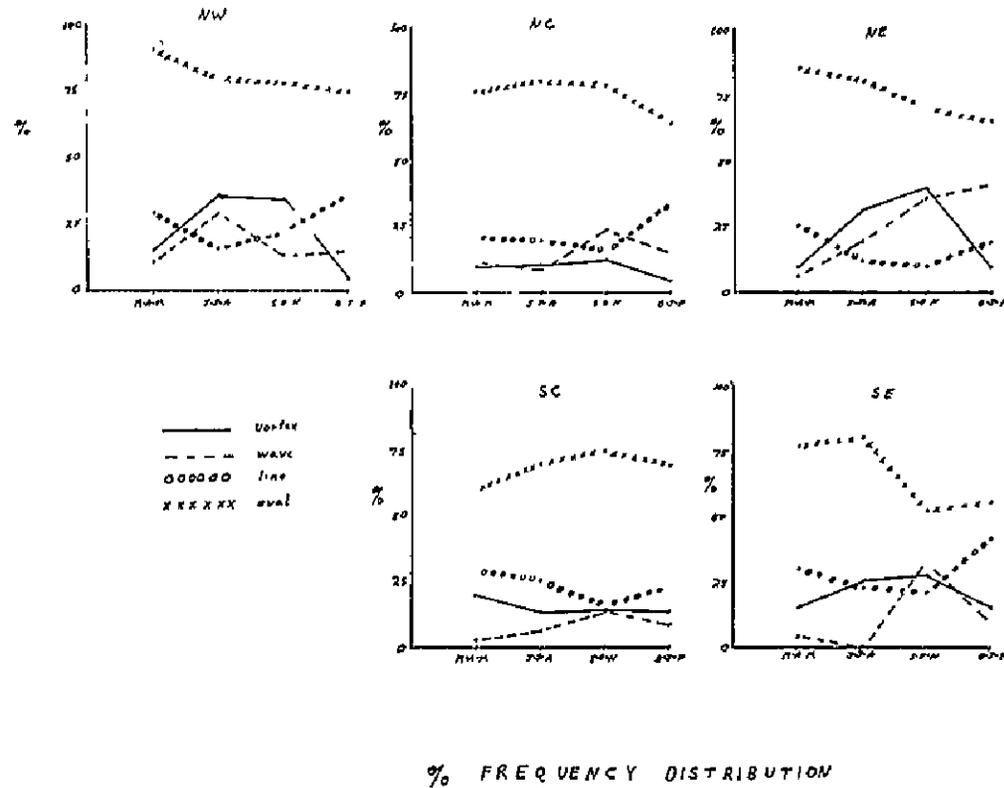
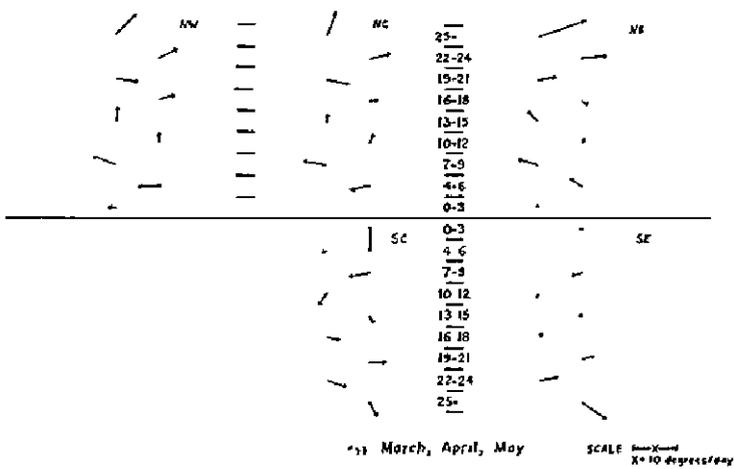
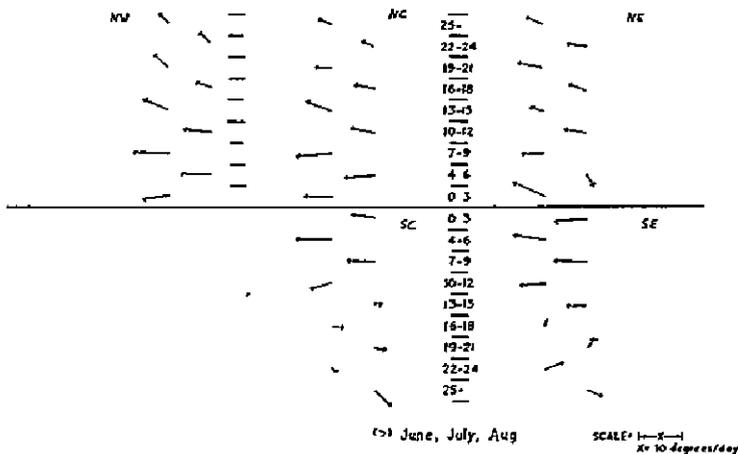


Fig. 9. The percent of cloud systems that reach a level of organization of oval, line, wave, and vortex for each season and region.

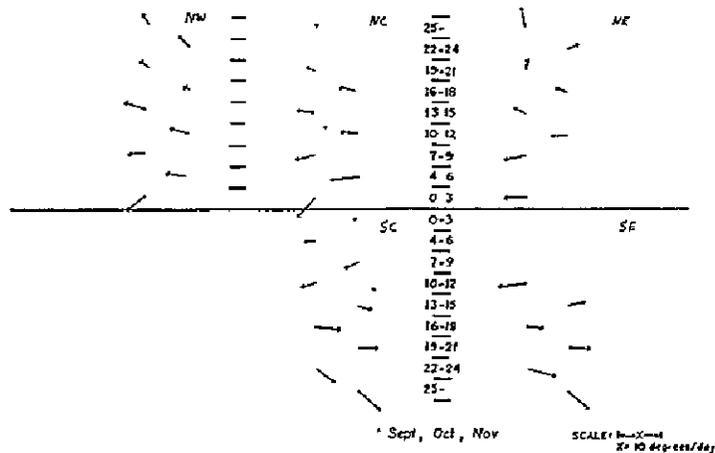


(a) March-April-May

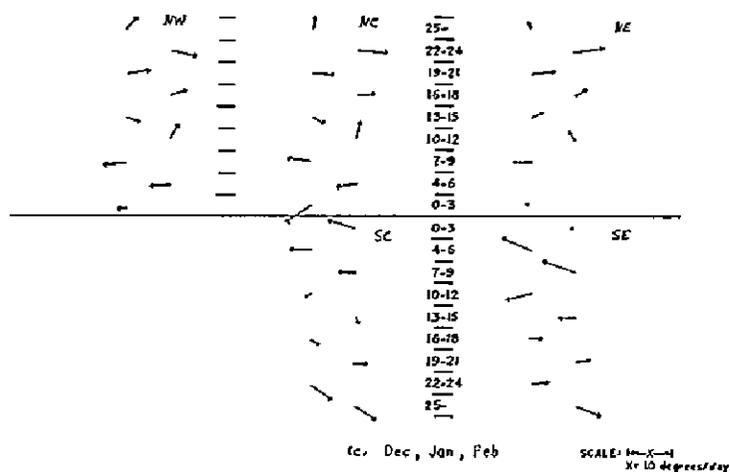
Figs. 10a, b, c, d. Mean displacement of cloud systems in three degree latitudinal strips by region. Arrow vector points in the direction the system was moving and represents the mean vector of all movements for that strip. 1 inch = 10 degrees per day.



(b) June-July-August



(c) September-October-November



(d) December-January-February

Wiederanders (1961) used a number of radiosonde observations from islands in the Pacific, and compiled mean wind direction and speed at five levels in the atmosphere for each month. His streamline analysis at the 500 mb and 200 mb level for July and January are shown in Figs. 11a, b and 12a, b, respectively, with the cloud census regions superimposed.

One can compare the streamline analyses with the mean movement of cloud systems for June - July - August and December - January - February. In Wiederanders July streamline analysis in the northern hemisphere, the easterly winds extend from the equator to 25°N in all five pressure levels, except in

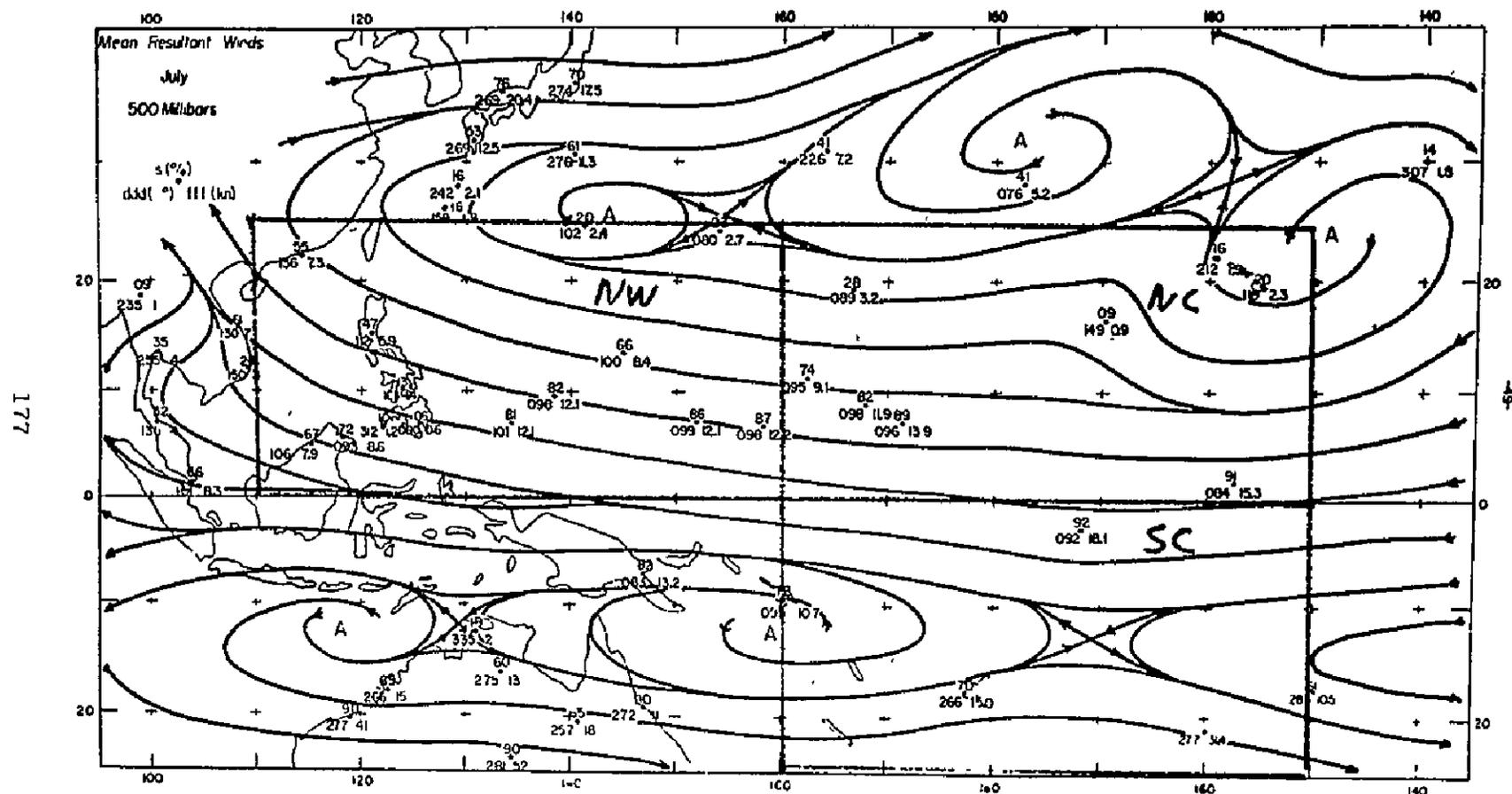


Fig. 11a. Streamlines for the central and western Pacific at 500 mb for July.

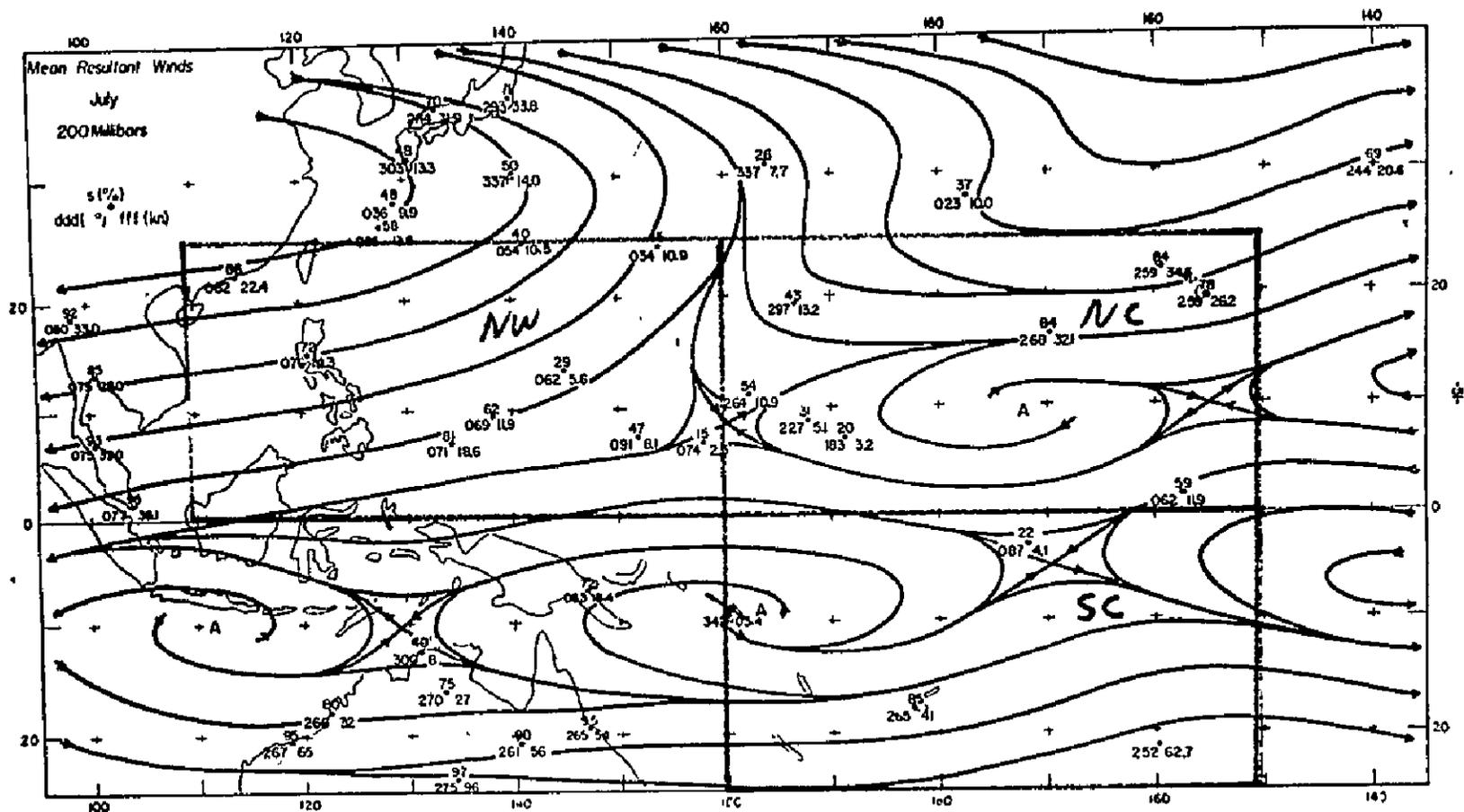


Fig. 11b. Streamlines for the central and western Pacific at 200 mb for July.

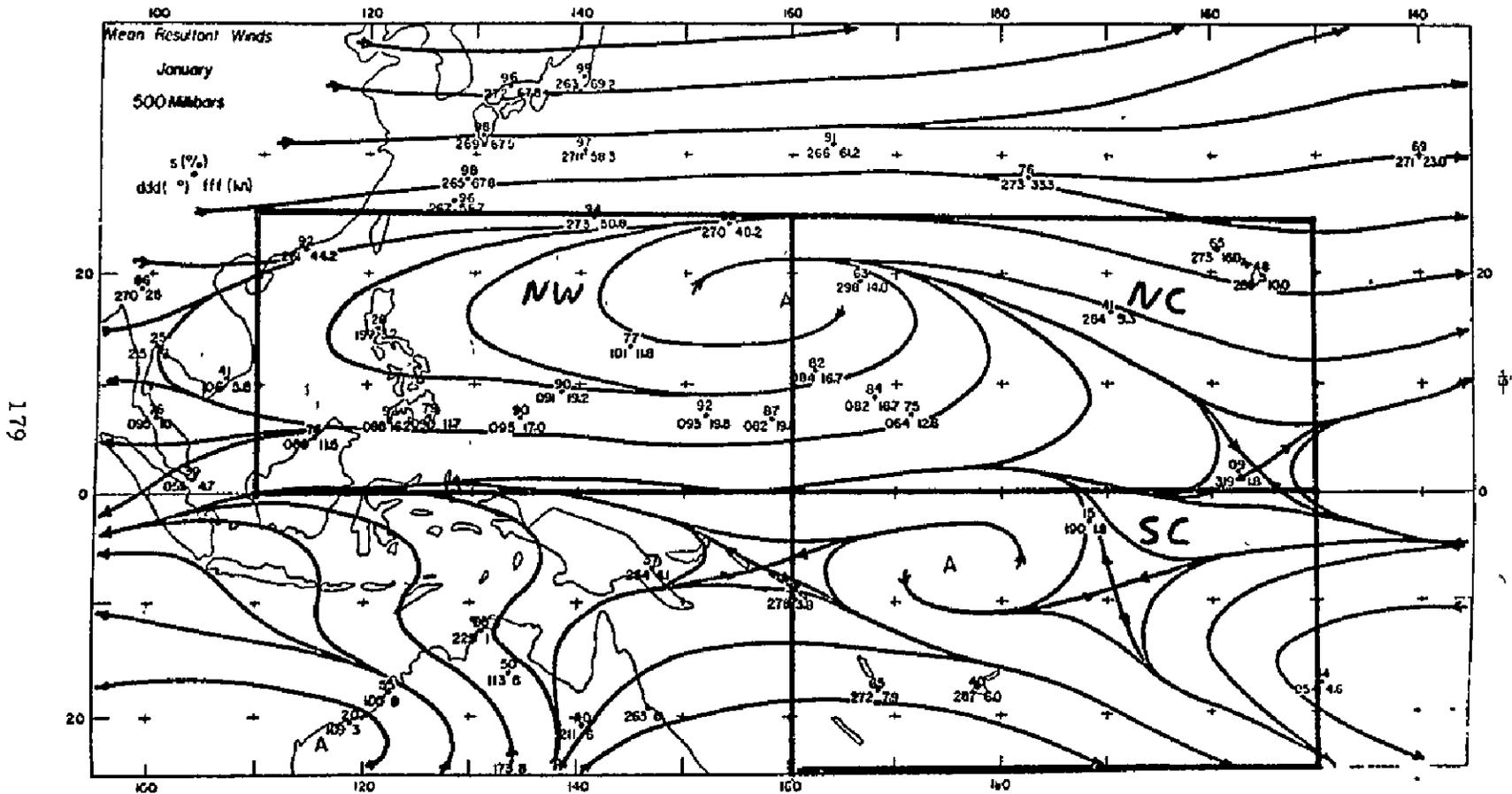


Fig. 12a. Streamlines for the central and western Pacific at 500 mb for January.

the northcentral region. In this region winds turn westerly at 300 mb and 200 mb for latitudes 10 - 25°N. Therefore, the mean movement of cloud systems is similar to the streamline flow in the northwest region, but is in disagreement with the flow in the upper troposphere in the northcentral region.

January streamline analysis is also in disagreement in the upper troposphere with the mean cloud movement in the northcentral region. From Wiederanders' analysis, westerly winds extend from 25°N to 10°N at 500 mb, and from 25°N to the equator at 300 mb and 200 mb. Therefore, the mean cloud movement which shows westward movement from latitudes 0° - 12°N, is consistent with the streamline flow at 500 mb but not at the upper troposphere. In the northwest region the mean cloud movement and streamline analysis are in agreement in the 500 mb and upper troposphere.

The mean wind speed, as calculated by Wiederanders, is greater than the mean movements of cloud systems at 300 mb and 200 mb levels. At the 700 and 500 mb regions mean wind speed is equal to or less than mean cloud movements.

5. Length of Life

In Fig. 13 the lifetime of cloud systems averaged over the entire year is plotted by region. For all regions and seasons vortices were the longest lived systems (6 days), ovals the shortest lived (2 days). Lifetimes varied most from region to region in case of lines (2 days northcentral to 7 days southcentral) and least for ovals (2 days in all regions).

Seasonal changes in lifetime (Fig. 14) were similar to regional variations. Lines varied most strongly from a minimum lifetime of just over 1 day (northcentral, Sept. - Oct. - Nov.) to a maximum lifetime of 11 days (southcentral, Sept. - Oct. - Nov.). The lifetime of vortices was also quite variable, especially in the northeast region. In all regions vortices lived longest during the summer months. Waves showed no consistent pattern of variation regionally; ovals were quite uniform.

One can compare the lifetime of reported hurricanes or typhoons in 1967 (from Mariners Weather Log) to vortex lifetime distribution as recorded in Fig. 14. In the northwest region in the summer and fall months the average vortex and typhoon lifetimes are about the same. However, in the northeast region in June - July - August the mean vortex lifetime is about twice as great as hurricane lifetimes. This means that a substantial number of the reported vortices have a level of organization lower than a hurricane. In September - October - November the average lifetimes of vortices and hurricanes are about equal.

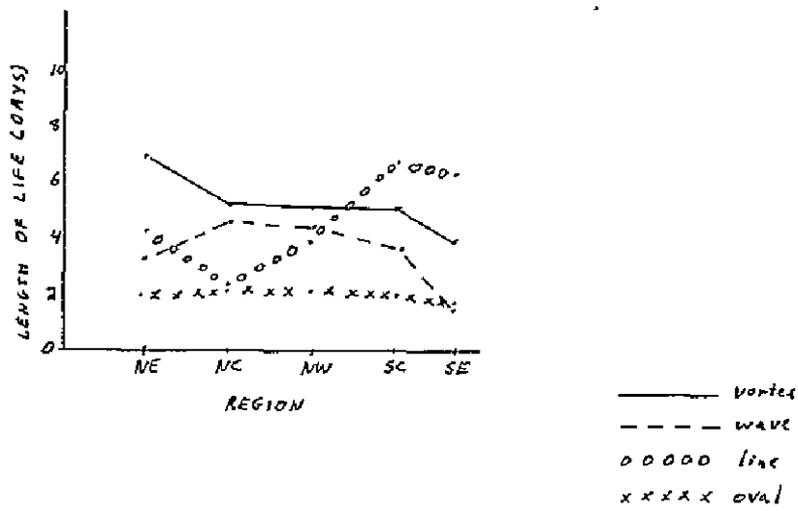


Fig. 13. Lifetime of cloud systems averaged over the entire year by region.

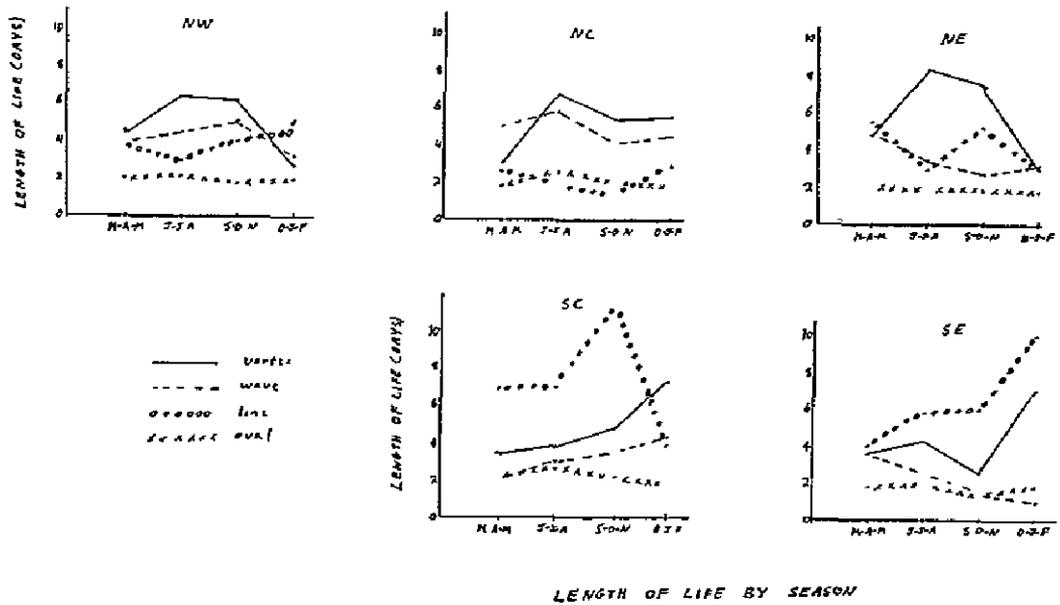


Fig. 14. Lifetime of cloud systems averaged per season and region.

6. Cloud Areas

In Fig. 15 the cloud area of systems averaged over the entire year is plotted by region. With the exception of lines only those systems wholly contained within the major boundary of the study area were included in the cloud averages.

It can be seen that systems tend to be smaller in the northeast and southeast regions than in the western regions of each hemisphere. Over the entire Pacific waves and lines were 50% larger in cloud area than ovals and vortices.

Seasonal variations in cloud area (Fig. 16) tend to be quite variable with ovals showing the only consistency from season to season. During summer and fall in the northern hemisphere systems tended to be smaller than in the winter months. In the southern hemisphere lines were remarkably large. Waves in the northeast region were also larger in the winter and spring seasons. One large and persistent vortex caused the unusually large increase in vortex area in the southcentral during December - January - February.

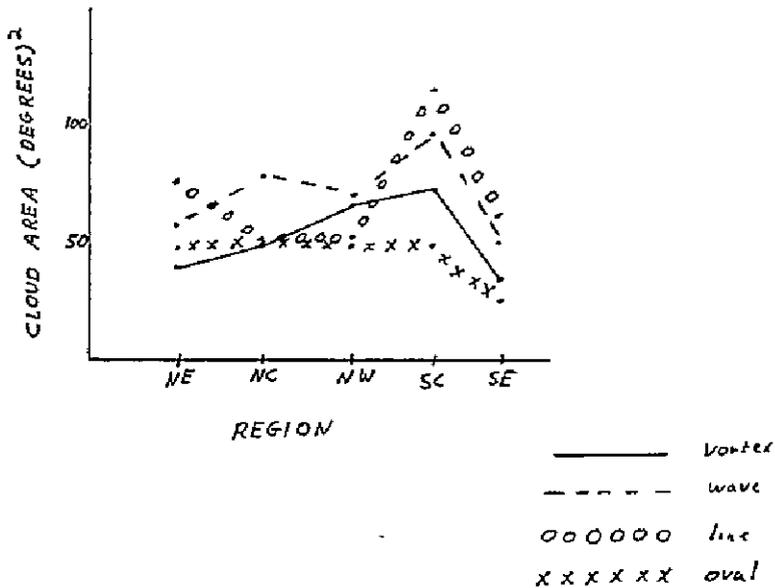
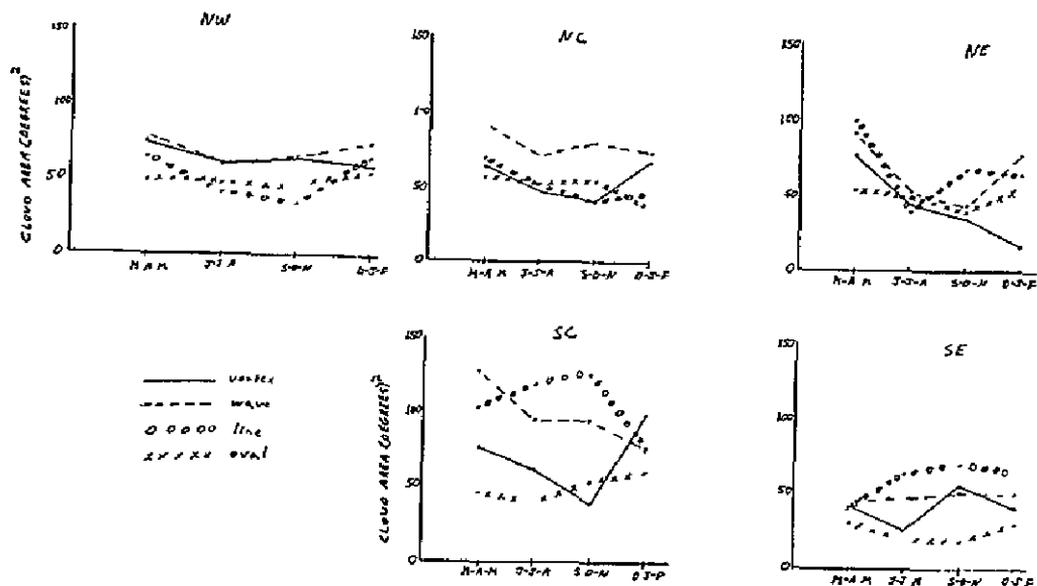


Fig. 15. Cloud area by cloud systems averaged over the entire year by region.



CLOUD AREA BY SEASON

Fig. 16. Cloud area of cloud systems averaged per season and region.

V. Conclusion

As a result of the cloud census we find that the northeast tropical Pacific Ocean contains an equally large percent of cloud systems that reach the organizational level of wave or vortex as in the northwest region. The northcentral Pacific region has a lower level of organization in the summer months than the two other regions in the northern hemisphere. There are considerable differences between the northern and southern hemisphere in the level of organization and the frequency of cloud systems.

VI. Bibliography

- Bolin, B., 1967. Report of the Study Conference in Stockholm, June 28 - July 11, 1967 ICSU/IUGG - Committee on Atmospheric Sciences, WMO - GARP.
- Gray, W. M., 1968. Global View of the Origin of Tropical Disturbances and Storms, Monthly Weather Review, Vol. 96, Number 10, pp. 669-700.
- Kornfield, J., and Hasler, F., 1969. A Photographic Summary of the Earth's Cloud Cover for the Year 1967, Journal of Applied Meteorology, pp. 687-700.

Mariners Weather Log, 1968, Vol. 12, Number 2, pp. 46 - 47.

Mariners Weather Log, 1968, Vol. 12, Number 3, pp. 77 - 78.

Martin, D. W., and Karst, O. J., 1970. A Census of Cloud Systems Over the Tropical Pacific, Department of Meteorology, University of Wisconsin, ESSA Annual Report 1968.

Wiederanders, C. J., 1961. Analysis of Monthly Mean Resultant Winds for Standard Pressure Levels Over the Pacific, Hawaii Institute of Geophysics, Report No. 13.

N71-11618

A CENSUS OF SYNOPTIC SCALE DISTURBANCES OVER THE CENTRAL AND
EASTERN PACIFIC DURING MARCH, 1967 - FEBRUARY, 1968

A. Staver¹, T. Vonder Haar, R. Cram and R. DeDecker

1. Introduction

Meteorological satellites are now providing a great deal of cloudiness information. This study classifies and counts the kinds of phenomena that form, move through and dissipate in defined regions of the tropics and mid-latitudes over the Pacific for the period of March, 1967 through February, 1968. The area studied corresponds to the usable ATS-I data. Its division into 14 nearly equal area regions is shown in Fig. 1. The once-a-day ESSA mercator and polar mosaics were also used. The initial criteria for identifying features was taken from Widger *et al.* (1965) and Simpson *et al.* (1968). It overlaps the study of Martin and Karst (1969) who studied the tropical features during 1967; however, the present study is less detailed, counts only large-scale features, and includes the mid-latitudes.

2. Census Methods and Definitions

The census taker's attention is focused upon the regions within one zone, such as the Northern Hemisphere Sub-Tropical zone which contains three regions. Each day of the month under study a census of synoptic-scale features in each region is recorded, numbered and thereby tracked. Features differentiated in subtropical zones included:

- a. Extratropical cyclones, clouds in a circular pattern with frontal clouds. The stage of development, developing, mature or dissipating is noted.
- b. Vorticity maxima, bright clouds indicating future development usually developing behind a low with no frontal association.
- c. Frontal waves, long band of clouds developing a warm and cold frontal configuration with no closely associated low initially.

¹Present affiliation: Northern Illinois University

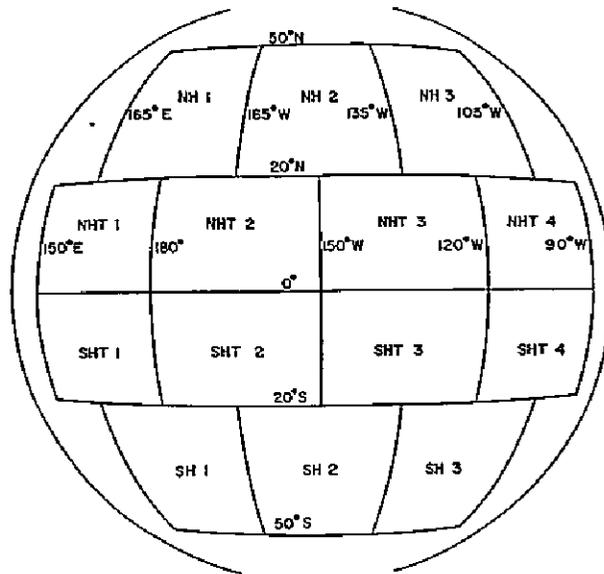


Fig. 1. Regions in the sub-tropics and tropics designed to be nearly equal areas in the usable portion of ATS-I photographs.

Features differentiated in tropical zones included:

- a. Migrating disturbance, bright clouds with no rotational motion moving with time.
- b. Non-migrating disturbance, similar; however, remaining stationary.
- c. Tropical wave, the "easterly" wave, a trough or cyclonic curvature maximum, forming an inverted V.
- d. Tropical depression, bright clouds with rotational motion.
- e. Typhoons, large bright clouds with strong rotational motion.

To test the objectivity of such a census, for some months and zones the same task was performed independently by another individual.

3. Comparison of Results Obtained by Two People Independently Reading the Same Data

a. Subtropics

1) Extratropical cyclones. For a particular month in the Northern Hemisphere one man recorded 9 lows, the other 10. There is perfect agreement on 8 of these, the others cut across or near the corner of the survey area. In the Southern Hemisphere, what one individual called a low was frequently called a vorticity maximum by the other man.

2) Vorticity Maxima. There is little correlation between two individuals trying to pick out small distinct vortices in the cold air west of major cyclones.

3) Frontal Waves. One individual recorded half as many frontal waves as the other. Succeeding photos were available to both individuals, one individual recorded only those waves which subsequently developed. If only the smaller number of frontal waves is considered, there is near perfect agreement between the two independent surveys.

In theory it would be nice to see a parent low and frontal system with a vorticity maximum approaching the front producing a developing frontal wave which subsequently becomes a new low with the old low dying. In this survey, if this process is considered one phenomenon the events recorded by two individuals show good agreement in the subtropics. Most discrepancies occur when one considers the breakdown into vorticity maximum, frontal wave and low; there is a lot of subjectivity in deciding which stage of development has been reached.

b. Tropics

There is little resemblance between phenomena recorded by two individuals. For example, one recorded 29 while the other recorded 46 separate phenomena for May in this sector of the Northern Hemisphere tropics. They are evidently recording patterns which are actually very transient. Nevertheless, there are significant findings:

1) There are 40 to 50 percent more tropical disturbances in the Northern Hemisphere than in the Southern Hemisphere in the studied area.

2) There is good agreement on disturbances which move from west to east. Some of these point toward the formation of tropical depressions some distance farther east after the cloudiness which was moving from west to east has nearly dissipated.

3) There is, of course, excellent agreement on tropical depressions and typhoons. It is noted that when a tropical depression forms, other disturbances in quite a large surrounding area die. When the depression reaches maturity, smaller disturbances reappear in the surrounding area. The formation of a depression therefore seems to entail the consolidation of vorticity from several small circulations.

4) Even if one chooses to discredit some of the disturbances as being the juxtaposition of individual cloud masses with wave motions producing the illusion of moving physical phenomena, the total number each month must be related to the number of "real" disturbances. Particularly when the results from one individual are studied, the difference in average duration of disturbances during each month and the number passing from one region to another must be considered significant.

4. Results Obtained in the Subtropics

Table 1 gives a broad scale view of activity in each subtropical region. Most notable are the large number of phenomena in both hemispheres during October, and the minimum number of phenomena in both hemispheres during July. The Southern Hemisphere displays a semiannual cycle in the number of phenomena with maxima during the equinoxes and minima at the solstices. A simplified view of the Northern Hemisphere numbers would indicate a maximum occurring from October through March and a minimum from April through September.

The average number of phenomena shown in the bottom line of Table 1 point out that in 6 of the 12 months, region 3 in the Northern Hemisphere contained many fewer phenomena than regions 1 or 2. A generalized explanation which will be supported by other tables indicates that region 2 exported as many disturbances through its northern boundary as were sent on eastward into region 3. This would indicate the presence of a ridge somewhere near the boundary between regions 2 and 3. The Southern Hemisphere had a nearly equal number of occurrences in all regions, the only exception being the small number in region 1 as compared to region 2 in May. The nearly equal number in all Southern Hemisphere regions leads to a conclusion of zonal flow as is usually expected.

The duration information is hampered somewhat by following these disturbances only in the sector under study (3 regions). Durations within the individual regions were available, with these and the 4.2 to 4.9 days indicated in Table 1, a good extrapolation would indicate a 6 day lifetime for average phenomena.

Table 1. Number of phenomena (total of lows, vorticity maxima and frontal waves) observed in each subtropical region and the average duration within the subtropical sector (3 regions).

Duration in days			Number of Phenomena			Month	Number of Phenomena			Duration in days		
NH1	NH2	NH3	NH1	NH2	NH3		SH1	SH2	SH3	SH1	SH2	SH3
3.8	4.1	4.8	12	14	10	Mar 67	12	13	8	3.5	4.1	3.5
3.7	4.1	3.4	10	9	7	Apr	11	14	9	3.7	4.6	4.4
4.9	4.8	2.7	9	10	3	May	4	10	9	6.2	5.5	6.1
6.1	6.1	4.9	8	8	8	June	7	8	6	4.6	4.8	4.0
4.7	3.7	3.0	9	7	1	July	3	6	4	6.0	5.0	5.5
5.2	5.7	4.5	12	10	4	Aug	10	10	8	5.9	7.2	5.1
3.8	3.6	3.8	12	9	4	Sep	8	10	7	4.6	3.8	4.1
4.6	5.1	5.4	16	16	10	Oct	16	19	12	4.3	4.7	3.8
4.1	4.3	4.0	13	15	10	Nov	15	13	13	4.3	5.5	5.2
4.3	3.9	3.8	11	11	4	Dec 67	9	7	9	3.3	5.0	3.7
4.6	4.6	6.1	16	18	7	Jan 68	7	9	8	4.9	4.4	3.9
3.8	4.5	4.2	15	13	9	Feb	7	10	7	4.3	4.6	5.1
4.4	4.5	4.2	12	12	6	Average	9	11	8	4.6	4.9	4.5

A rather surprising statistic is the fact that only 30 of the 360 phenomena traversed a region in the Northern Hemisphere, and even less, 17 of 338 entered the western boundary and left via the eastern boundary in the Southern Hemisphere. None went through 2 regions in the Southern Hemisphere while 6 did accomplish it in the Northern Hemisphere. Carrying this to its conclusion, none in either hemisphere ever entered the west boundary of region 1 and exited from the east boundary of region 3. Frontal zones may propagate downstream, but cyclogenetic baroclinic portions of the atmosphere seem to form over a particular area of the earth's surface. The lows formed by the cyclogenesis evidently travel from western portions of this area to maturity at the eastern and after which they decelerate and dissipate. To traverse the 30° of longitude of each region in this study required an average of 3 days in the Northern Hemisphere compared to 4 days in the Southern Hemisphere. From Table 1 it should be noted that the average lifetime was 4.5 days, while the figures stated above indicated only 7% managed to traverse a region requiring 3.5 days travel time. The majority of phenomena must be quite slow moving; in other words, fast moving lows manage to traverse regions. There is no apparent difference in duration between speedy and nearly stationary cloud systems.

Table 2. Number of phenomena moving west to east at region boundaries and moving poleward through the poleward region boundary.

	Number moving west to east at region boundary						Month	Number moving poleward through poleward region boundary							
	NH1	NH2	NH3	SH1	SH2	SH3		NH1	NH2	NH3	SH1	SH2	SH3		
192	2	5	4	1	1	5	3	1	Mar 67	2	2	3	2	2	5
	3	1	2	0	1	4	3	6	Apr	7	1	2	2	5	2
	2	5	1	0	0	3	6	5	May	0	6	1	0	0	1
	2	3	1	2	1	5	2	1	June	2	6	0	0	0	0
	4	1	0	0	1	0	3	0	July	4	2	0	0	0	0
	7	4	0	0	6	3	4	5	Aug	2	6	0	0	0	0
	2	6	0	0	1	5	0	0	Sep	2	5	1	1	3	1
	11	9	7	1	5	10	7	5	Oct	2	1	6	2	4	2
	8	5	7	1	6	7	5	5	Nov	3	1	0	2	5	1
	4	3	2	0	2	1	2	3	Dec 67	0	2	0	4	4	3
	8	8	5	0	1	2	2	2	Jan 68	1	1	0	5	3	2
	5	6	5	1	2	2	4	1	Feb	1	3	3	4	2	1
	58	57	34	6	27	47	41	34	Total	26	36	16	22	28	18
	5	5	3	1	2	4	3	3	Average	2	3	1	2	2	2

Table 2 contains the number of phenomena moving west to east at each region's boundary and the number leaving the region via the poleward boundary. Again the average number moving west to east indicate a more uniform flow in the Southern Hemisphere. The increase from 2 at the western boundary to the Southern Hemisphere's region 1 to 4 at the eastern edge is due to cyclogenesis as will be shown by Table 3. Table 3 will also explain the difference in average numbers crossing east-west boundaries in the Northern Hemisphere as being due to formation of cloud systems in region 1 as opposed to a sink in region 3 combined with a large number going toward the pole in region 2. It remains to be checked whether a large number going across the 50° latitude line toward the pole indicates a trough somewhere near the western boundary of that region. If it does, then the Northern Hemisphere should have a trough near 165E in April, 1967, one near 165W in May, June, August and September and a trough near 135W in October. If these prove to be reasonably accurate, the following mean trough positions can be implied in the Southern Hemisphere: near 135W in March, 1967, near 165W in April and November, and near 165E in January, 1968. In both hemispheres more phenomena go poleward across the 50° latitude line during their respective summers. There is a very definite minimum in the Northern Hemisphere in December and January and in the Southern Hemisphere from May through August.

Table 3 shows the statistics on formation and dissipation in each region as well as their algebraic subtraction to determine whether a particular region in each month is a source or sink of cloud systems. In this subtropical sector there is a minimum of cyclogenesis in both hemispheres during the month of July. There is a marked lack of cyclogenesis in region 3 of the Northern Hemisphere in all except March, April and June. If one wanted to study cyclogenesis, it would be best accomplished where the ratio of cyclogenesis to the total number of phenomena is large. This ratio is largest in December through April in both hemispheres, with the exception that region 3 in the Northern Hemisphere is a poor region to study cyclogenesis except in the months noted above.

Figure 2 is an overall summary of phenomena entering and leaving the subtropical sectors as well as illustrating that formation and dissipation far overshadow the number crossing the boundaries. The Southern Hemisphere sector is a source—formation exceeds dissipation by 56—whereas the Northern Hemisphere sector shows a net dissipation of 5. Table 3 showed more detail indicating that in the Northern Hemisphere region 3 is a sink and region 1 a source, while all regions in the Southern Hemisphere are sources, particularly region 1. The 9 existing and 7 existing shown in Fig. 2 are an admitted mathematical uncertainty created by beginning and end of the month discontinuities in the census taking. The census taker was not allowed to study what he had recorded from chronological months, the randomized procedure produced 9 more phenomena recorded on the firsts of the months as compared to the last day of the preceding months in the Northern Hemisphere, for example. Except for

Table 3. Formation and dissipation of phenomena in each subtropical region.

	Number formed in each region						Number dissipated in each region						Month	Source (+) or Sink (-) of Phenomena					
	NH1	NH2	NH3	SH1	SH2	SH3	NH1	NH2	NH3	SH1	SH2	SH3		NH1	NH2	NH3	SH1	SH2	SH3
	10	7	5	8	7	5	4	8	6	5	5	2		Mar 67	6	-1	-1	3	2
4	7	5	7	9	5	1	6	5	3	3	1	Apr	3	1	0	4	6	4	
5	3	1	4	6	3	4	2	2	1	3	2	May	1	1	-1	3	3	1	
4	4	5	6	3	3	2	1	6	2	5	4	June	2	3	-1	4	-2	-1	
5	2	0	1	5	1	3	5	1	1	3	3	July	2	-3	-1	0	2	-2	
1	5	0	3	5	4	5	4	3	6	3	3	Aug	-4	1	-3	-3	2	1	
6	3	2	5	4	5	2	4	3	1	6	6	Sep	4	-1	-1	4	-2	-1	
4	5	1	7	4	5	4	6	3	3	8	6	Oct	0	-1	-2	4	-4	-1	
4	8	3	9	4	6	5	6	8	4	3	6	Nov	-1	2	-5	5	1	0	
6	7	2	6	3	7	6	6	4	4	1	2	Dec 67	0	1	-2	2	2	5	
7	10	2	5	6	4	6	10	7	0	4	4	Jan 68	1	0	-5	5	2	0	
8	6	2	3	6	3	7	5	4	1	3	4	Feb	1	1	-2	2	3	-1	
64	67	28	64	62	51	49	63	52	31	47	43	Total	15	4	-24	33	15	8	
5	6	2	5	5	4	4	5	4	3	4	4	Average	1	0	-2	3	1	1	

194

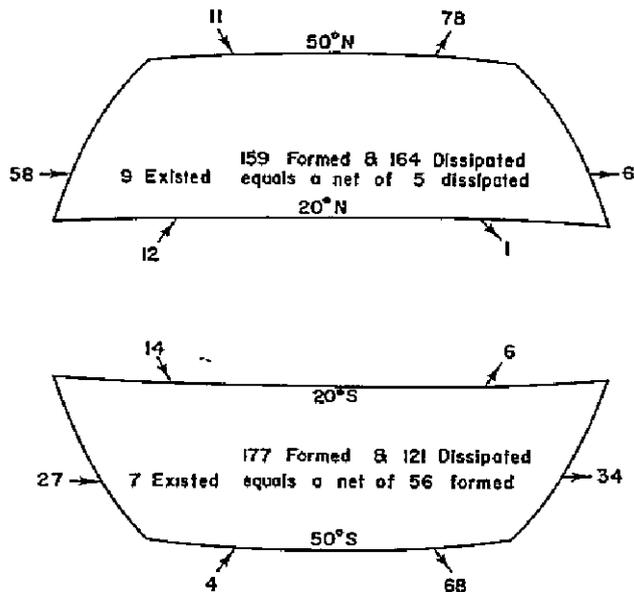


Fig: 2. Phenomena entering and leaving the subtropical sectors as well as number formed and dissipated.

the significant number going poleward at the 50° latitude line, very few phenomena were observed to cross the northern and southern boundaries.

5. Results obtained in the tropics

Tables 4, 5 and 6 present the census of tropical phenomena by region and month. They are presented in increasing order of meteorological organization: migratory disturbances (MD), tropical waves (W), tropical depressions (TD), and typhoons (TYP). The maximum number of MDs are found in the western region of each hemisphere—that is, the portion of the hemisphere included in this study. The decrease in number of occurrences in the eastern regions is much more marked in the Southern Hemisphere. In fact, in the Southern Hemisphere's region 4 virtually none are found except in March through May. The number of tropical waves is a maximum in the Northern Hemisphere's region 3; here are very few in the Southern Hemisphere. In the two types mentioned, MDs and Ws, there is little variation in frequency month to month with a slight tendency for a minimum near the solstices. Tropical depressions and typhoons are absent during each hemisphere's respective winter. The total number is far smaller in the Southern Hemisphere. The Southern Hemisphere maximum is in its western region with none occurring in its two eastern regions. The Northern Hemisphere maximum is in its eastern region, which is exactly opposite that of MDs. There are virtually no TDs or typhoons in Northern Hemisphere region 2.

Table 4. Number of migratory disturbances and tropical waves

Number of Migratory Disturbances				Month				Number of Tropical Waves								
NHT1	2	3	4	SHT1	2	3	4		NHT1	2	3	4	SHT1	2	3	4
12	10	10	6	13	14	10	5	Mar 67	1	2	4	2	1	2	2	1
10	12	15	9	11	12	6	7	Apr	0	2	5	1	0	2	1	0
13	13	10	11	10	7	6	7	May	2	3	4	1	2	0	0	0
15	9	11	11	12	7	4	2	June	0	0	0	1	0	0	0	1
9	6	7	6	3	1	0	0	July	1	1	3	2	0	0	0	0
13	10	10	9	12	5	3	1	August	2	1	1	2	0	0	0	0
14	8	8	8	12	8	3	0	Sep	0	2	1	0	0	0	0	0
14	13	7	8	11	6	2	0	Oct	0	0	2	0	0	0	0	0
7	12	8	5	11	7	1	0	Nov	0	1	3	0	0	0	0	0
7	5	10	7	11	9	3	0	Dec 67	0	2	1	0	0	0	0	0
12	11	12	8	10	7	3	0	Jan 68	0	0	0	0	0	0	0	0
7	5	3	4	8	4	1	0	Feb	0	0	0	1	0	0	0	0
133	114	111	92	124	87	42	22	Totals	6	14	24	10	3	4	3	2

Table 5. Number of Tropical Depressions

Total number in Tropical Sector (N Hemi)	Number in Each Region			Month	Number in Each Region			Total number in Tropical Sector (S Hemi)		
	NHT1	2	3		4	SHT1	2		3	4
1	1	0	0	0	Mar 67	0	0	0	0	1
0	0	0	0	0	Apr	0	1	0	0	1
1	0	0	1	1	May	0	1	0	0	1
4	0	0	0	4	June	0	0	0	0	0
6	1	1	4	3	July	0	0	0	0	0
7	1	0	3	5	Aug	0	0	0	0	0
7	2	0	1	4	Sep	0	0	0	0	0
10	1	1	3	6	Oct	0	0	0	0	0
3	3	0	0	0	Nov	4	0	0	0	4
0	0	0	0	0	Dec 67	0	1	0	0	1
0	0	0	0	0	Jan 68	1	0	0	0	1
0	0	0	0	0	Feb	2	1	0	0	3
39	7	2	12	23	Totals	8	4	0	0	12

Table 6. Number of Typhoons

Total number in Tropical Sector (N Hemi)	Number in Each Region				Month	Number in Each Region				Total number in Tropical Sector (S Hemi)
	NHT1	2	3	4		SHT1	2	3	4	
1	0	0	0	0	Mar 67	1	0	0	0	1
0	0	0	0	0	Apr	0	1	0	0	1
1	0	0	1	0	May	0	0	0	0	0
2	0	0	1	2	June	0	0	0	0	0
2	0	0	2	0	July	0	0	0	0	0
3	1	0	1	2	Aug	0	0	0	0	0
6	2	1	4	3	Sep	0	0	0	0	0
5	1	0	1	3	Oct	0	0	0	0	0
2	2	0	0	0	Nov	2	0	0	0	2
0	0	0	0	0	Dec 67	0	1	0	0	1
0	0	0	0	0	Jan 68	1	0	0	0	1
0	0	0	0	0	Feb	0	1	0	0	1
22	6	1	10	10	Totals	4	3	0	0	7

In Tables 5 and 6 the number in each region can be totaled for the four regions and compared to the total number in the sector (already tabulated), to ascertain how many tropical storms moved from one region to another. In the Southern Hemisphere none managed to cross a north-south boundary between two regions in the tropics. Since none occurred in regions 3 and 4, this merely means none of the 7 storms (4 TDs and 3 typhoons) went westward through the 180° longitude line. In the Northern Hemisphere there is some movement indicated in June through October.

The lifetimes of these tropical phenomena as observed within the confines of the tropical sector studied averaged as follows (in days):

	Migratory Disturbance	Tropical Wave	Tropical Depression	Typhoon
Northern Hemisphere	5.0	3.3	2.8	4.1
Southern Hemisphere	5.1	3.3	2.7	4.4

Most striking is the similarity in duration when comparing the two hemispheres' figures for each type of cloud system. The majority of migratory disturbances dissipate, some move out of the sector under study and a few develop into the more organized types of phenomena. As shown in Tables 4, 5 and 6, each stage of organization has a smaller number of occurrences. The central stages, Ws and TDs, require a shorter time before they either dissipate or develop, which indicates they are not a dynamically stable configuration. Most of the

typhoons moved out of the sector under study, so their total lifetime is even longer.

6. Interchange between Hemispheres and between the Tropics and Subtropics of each Hemisphere

The only cloud systems, on a scale considered in this study, which crossed the equator were classified as migratory disturbances. Seven went from the Northern Hemisphere to the Southern all in region 1 in March, April, June and December. Three went the other way in region 1 in April, June and December. Another three went from the Southern to the Northern Hemisphere in region 2 during April, July and December. The compensating movements toward the north and south across the equator in the same months seem like more than a coincidence.

Table 7 shows the number of migratory disturbances going from the tropics to the subtropics and conversely, cloud systems coming from the subtropics into the tropical sectors surveyed. The numbers here should be about four-thirds times the numbers shown in Fig. 2, since the poleward boundary of the tropics was used here as compared to the shorter equatorward boundary of the subtropical sectors. The two hemispheres are remarkably similar with the majority of tropical disturbances going poleward in the western regions of the sector studied. Cloud systems coming into the tropics from the subtropics are infrequent that no generalizations can be surmised. As stated previously, only those that tracked from west to east seemed to point to development of tropical depressions. No general correlation exists between these occurrences and the number of TDs found in these regions.

Table 8 shows the regions and months where more highly organized tropical systems went into the subtropics. Of the seven typhoons in the Southern Hemisphere, five went directly into the subtropics from the region in which they formed. Even the other two weakened to tropical depressions, and one of those went into the subtropics from its original region. Both hemispheres display the familiar late summer, early autumn maximum in their respective hemispheres. Looking at the west to east distribution, the Southern Hemisphere has a maximum of these "storm recurvatures" in its westerly regions where Tables 5 and 6 showed the total number of organized storms was a maximum. Recurvatures in the Northern Hemisphere also correspond to the total number of storms. Region 2 had only 3 storms; none recurved.

Table 7. Migratory disturbances going poleward and equatorward through the 20° latitude line in each hemisphere.

Migratory Disturbances going into the subtropics								Month	Migratory Disturbances coming from the subtropics							
NHT1	2	3	4	SHT1	2	3	4		NHT1	2	3	4	SHT1	2	3	4
1	4	0	0	3	2	2	0	Mar 67	1	0	0	0	0	2	1	0
2	1	2	0	2	2	0	0	Apr	0	0	1	0	0	1	1	0
1	1	1	0	1	2	1	0	May	0	1	1	1	0	0	0	0
0	0	0	0	3	0	0	0	June	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	July	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	Aug	2	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	Sep	0	0	0	1	0	1	0	0
3	0	1	0	1	0	0	0	Oct	0	0	0	0	0	0	0	0
0	3	1	0	0	0	0	0	Nov	0	0	0	0	0	0	0	0
1	0	1	1	1	0	1	0	Dec 67	0	1	0	0	2	1	1	0
0	0	1	1	4	1	0	0	Jan 68	0	0	0	2	1	0	0	0
0	0	0	0	0	0	0	0	Feb	0	0	0	0	0	0	0	0
15	10	7	2	16	7	4	0	Totals	3	2	2	4	3	6	3	0

Table 8. Regions where Tropical Waves (W), Tropical Depressions (TD), and Typhoons (TYP) went poleward into the Subtropics

NHT1	NHT2	NHT3	NHT4	Month	SHT1	SHT2	SHT3	SHT4
				Mar 67	TYP	W		
		W		Apr		TYP		
				May	2Ws	TD		
		TYP	TYP	June				
TD		TD, TYP	TD	July				
		2TD, TYP	TD, TYP	Aug				
TD, 2TYP		3TYP		Sep				
			2TYP	Oct				
				Nov	TD			
		W		Dec 67		TYP		
				Jan 68	TYP			
				Feb	TD	TYP		

7. Summary and Conclusions

Census taking of meteorological phenomena from satellite cloud photographs leads to a certain amount of ambiguity when different individuals attempt to interpret the same basic data. Most of the difficulty is caused by rather disorganized cloud clusters which occur both in the tropics and subtropics. With this reservation several tables and discussion of number of occurrences and duration of various phenomena have been presented. A reader would be advised to select that information which applies to the problem he is considering. For tropical studies an area the size of a region as used in this study would suffice since most storms would intensify, dissipate or move into the subtropics from one region. For studying typhoons, Northern Hemisphere region 4 would be the most suitable of those regions included in this study. September would be the best month with a high number of occurrences against a background of only a few migratory disturbances. Tropical "Easterly" Waves would best be studied in Northern Hemisphere region 3 in April. Movement of cloud clusters across the equator in both directions as occurred between the regions labeled as 1 in April should be studied further. Northern Hemisphere region 2 and Southern Hemisphere regions 3 and 4 notably lacked phenomena. Turning the discussion to the subtropics, there Northern Hemisphere region 3 had a smaller number of occurrences. It remains to be determined whether a sizable export of systems toward the pole and/or a source of cyclogenesis within a subtropical region denote a long-wave trough near the western boundary of the region. The longer-lived phenomena were slow moving; this combined with an average lifetime of 6 days helps explain the surprisingly small number of phenomena which traversed an entire region. Formation and dissipation followed by reformation in a newly created cyclogenetic area seems to be the rocky pathway to be overcome by meteorological forecasters.

8. References

- Martin, D. and Karst, O., 1969: Cloud Census to Pacific Tropical Storms. ESSA Annual Report, University of Wisconsin, Madison.
- Simpson, R., et al., 1968: Atlantic Tropical Disturbances, 1967, Monthly Weather Review, pp. 251-259.
- Widger, et al., 1965: Practical Interpretation of Meteorological Satellite Data. U. S. Air Force Air Weather Service Technical Report No. 185, pp. 380.

Supported under Contract NAS5-11542.

N71-11619

POSSIBILITIES FOR QUANTITATIVE RADIANCE MEASUREMENTS IN THE
450-650 nm REGION FROM THE ATS-1 SATELLITE

Susan K. Peekna, Robert J. Parent and Thomas H. Vonder Haar

LIST OF SYMBOLS

A_C	: Area camera entrance aperture, m^2
A_S	: Area of source radiation, m^2
$A_{S\perp}$: Source area projected perpendicular to r_S , m^2
$H'(\delta)$: Effective irradiance on A_S from a zenith angle δ
m	: Optical air mass
$N_{\lambda S}$: Spectral radiance from a source, $watts\ m^{-2}\ ster^{-1}nm^{-1}$
$N_S'(\gamma)$: Effective radiance from A_S in direction γ , $watts\ m^{-2}\ ster^{-1}$
$N_R'(\delta)$: Effective radiance on A_S from direction δ , $watts\ m^{-2}\ ster^{-1}$
P'	: Effective power input to camera, watts
$P_{\lambda r}$: Power of wavelength λ received at camera entrance, $watts\ nm^{-1}$
r_S	: Distance from camera to A_S , m
r'	: Effective directional reflectance
r	: Total directional reflectance of Kodak white paper
R_λ	: Normalized spectral transmission of the camera optics
V_0	: Voltage output from the camera system, millivolts, volts
ϕ	: Azimuth angle in plane of A_S
γ	: Zenith angle of satellite from A_S , degrees, radians
δ	: Zenith angle of sun from A_S , degrees, radians
λ	: Wavelength, millimicrons (nanometers)
ρ'	: Effective bi-directional reflectance ($ster^{-1}$)
α	: Angular field of view of camera, radians
ω_C	: Solid angle field of view of camera, steradians
ω_S	: Solid angle subtended by the camera entrance aperture at the source
$d\Omega$: Elemental solid angle
Ω_\odot	: Solid angle subtended by the sun at A_S

1. Introduction

The Spin-Scan Camera experiment flown on ATS-I consists essentially of a photomultiplier tube and associated optics that respond to variations in a selected portion of solar energy reflected and scattered from the earth and its atmosphere. Other papers discuss the details of camera construction (Thomsen, et al. (1968)), and the pre-launch calibration (Suomi and Parent (1968)). They show that the spectral bandpass of the camera is from 450 to 650 m μ .

The purpose of this paper is to develop the equations that could be used in prelaunch calibration of satellite cameras such as that flown on ATS-I. We have used the prelaunch calibration that was done on ATS-I only as an example of how such a calibration might be performed, because these particular measurements were obtained under less than optimum conditions and were intended merely to make necessary prelaunch adjustments to the camera gain level setting. An absolute calibration of ATS-I cloud camera has been carried out after launch and is discussed by Hanson and Suomi (1968).

Effective radiance values can be used to investigate the radiation budget of the earth-atmosphere system and, together with surface observations, the absorption of solar energy in the atmosphere. They can also provide new information about the spectral and bi-directional reflectance characteristics of different surfaces. Variations of cloud thickness can be inferred from radiance values and the type and amount of cloudiness over a region can be related to the reflected energy. Among the other possible uses of these quantitative values are studies of the diurnal variation of cloudiness, background brightness of various regions and work oriented toward passive thermal control of spacecraft.

The advantages of applying ATS-I data to these problems stem from the high spatial resolution and large dynamic range of the sensors and the nearly continuous time sampling of the experiment. For many studies, only relative brightness measurements are required. This paper will present information about this type of application of the ATS-I data and will also discuss the procedure for obtaining absolute values of reflected spectral radiance. At the present time the most accessible quantitative values of ATS-I measurements are stored in digital form on magnetic tape. Thus, the following sections will consider the necessary steps required to convert these digital count values to the effective radiance of the source as viewed by the camera.

2. Total Power Input to the Camera in Terms of Earth Radiance

The total power incident on the camera entrance aperture within the field of view of the camera can be related to the radiance of the source. The geometry for this calculation is shown in Fig. 1, which depicts the instantaneous view of the camera of an area on the surface of the earth. Assuming the field of view is a cone of angle α with a uniform response (i. e., the response is

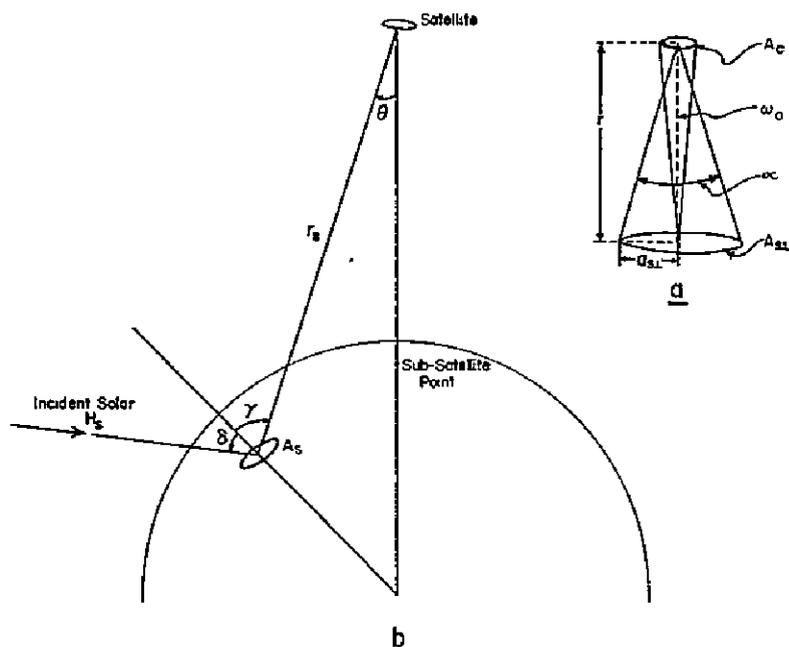


Fig. 1. Satellite-earth geometry

uniform within the cone of view and zero elsewhere), the spectral power $P_{\lambda r}$ received is from an earth A_S on the earth defined by the intersection of the cone and the earth. $P_{\lambda S}$ can be calculated under these conditions in terms of the radiance of A_S , $N_{\lambda S}$, following the definition of N_{λ} used by Nicodemus.* Assuming the radiating surface to be the top of the atmosphere and thus neglecting atmospheric absorption, we have,

$$P_{\lambda r} = A_S \omega_S N_{\lambda S}(\gamma) \cos \gamma \text{ [watts nm}^{-1}\text{]} \quad (1)$$

Where ω_S is the solid angle subtended at the center of A_S by the camera entrance aperture A_C . Using a small angle approximation,

$$\omega_S = \frac{A_C}{r_S^2} \text{ [sr.]} \quad (2)$$

Now let A_{sl} be the source area projected perpendicular to r_{sl} . Then

* Radiance is the radiant power (P) per unit solid angle (Ω) in the direction of a ray per unit projected area ($A \cos \theta$) perpendicular to the ray.

$$A_{SL} = A_S \cos \gamma$$

$$A_S = \frac{A_{SL}}{\cos \gamma} \quad (3)$$

The area A_{SL} is a circle, and its radius a_{SL} is given by small angle approximation, when α is expressed in radians, as,

$$a_{SL} = r_S \frac{\alpha}{2} \quad (4)$$

Substituting,

$$A_{SL} = \pi a_{SL}^2 = \frac{\pi r_S^2 \alpha^2}{4} \quad (5)$$

Then:

$$A_S = \frac{\pi r_S^2 \alpha^2}{4 \cos \gamma} \quad (6)$$

Substituting (6) and (2) into (1),

$$P_{\lambda r} = N_{s\lambda}(\gamma) A_C \pi \frac{\alpha^2}{4} \quad (7)$$

which is the desired relationship between $P_{\lambda r}$ and $N_{\lambda s}(\gamma)$.

3. Radiant Input Measured by the Camera

In order to interpret the output of the camera as a measurement of radiant input, it is important to note that it is not the total power input that is being measured, but only that part which lies within the spectral band width (450 - 650 nm) of the ATS-I camera. With this in mind, the "effective" power input, P' , to the camera is defined as:

$$P' = \int_{\lambda_1}^{\lambda_2} P_{\lambda r} R_{\lambda} d\lambda \quad (8)$$

where R_{λ} is the normalized spectral transmission of the camera optics, which includes the effects of the mirrors, lens, filter and photocathode surface, and λ_1 and λ_2 are the limits of the spectral bandwidth of the camera (Fig. 2).

A thorough examination of the camera response to a known source is essential in providing a reliable relationship between P' and camera output. This information will be provided by the inflight calibration. Substituting (7) into (8) gives the effective power received from the source as:

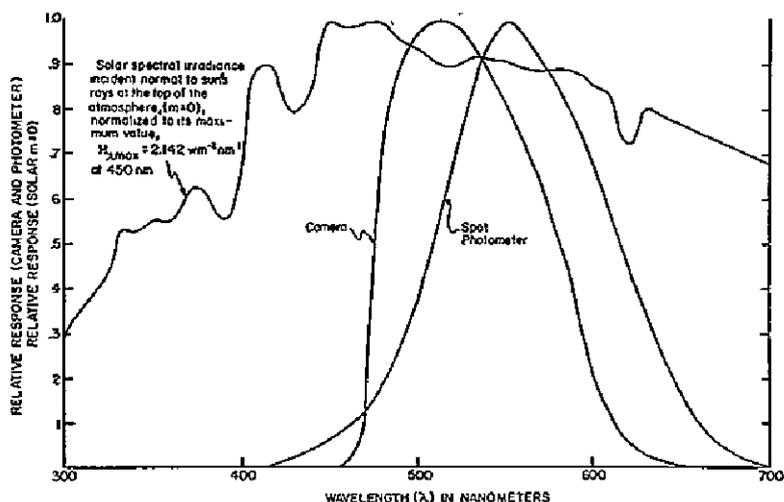


Fig. 2. Normalized spectral response of camera and photometer; normalized spectral solar irradiance on a surface perpendicular to the sun's rays for $m = 0$ at mean earth-sun distance.

$$P_r' = \frac{A_c \pi \alpha^2}{4} \int_{\lambda_1}^{\lambda_2} N_{\lambda S}(\gamma) P_\lambda d\lambda \quad (9)$$

The integral in this equation can be defined as the effective radiance from the source A_S in the direction γ , or $N_S'(\gamma)$. Using this definition,

$$P_r' = \frac{A_c \pi \alpha^2}{4} N_S'(\gamma) \quad (10)$$

Since for small angles, $\omega_c = \frac{\pi \alpha^2}{4}$, Eq. (10) can be expressed as:

$$P_r' = A_S \omega_c N_S'(\gamma) \quad (11)$$

This last equation is true in general only under the condition stated, i. e., the field of view is a cone with a uniform response, and the source viewed is radiating uniformly over the area A_S .

4. Bi-Directional Reflectance Measurements

Additional information can be derived from the ATS measurements of $N_S'(\gamma)$ when they are used to infer the effective bi-directional reflectance $\rho' = \rho(\delta, \phi_1, \gamma, \phi_2)$ of the region viewed by the camera (Fig. 3). The general

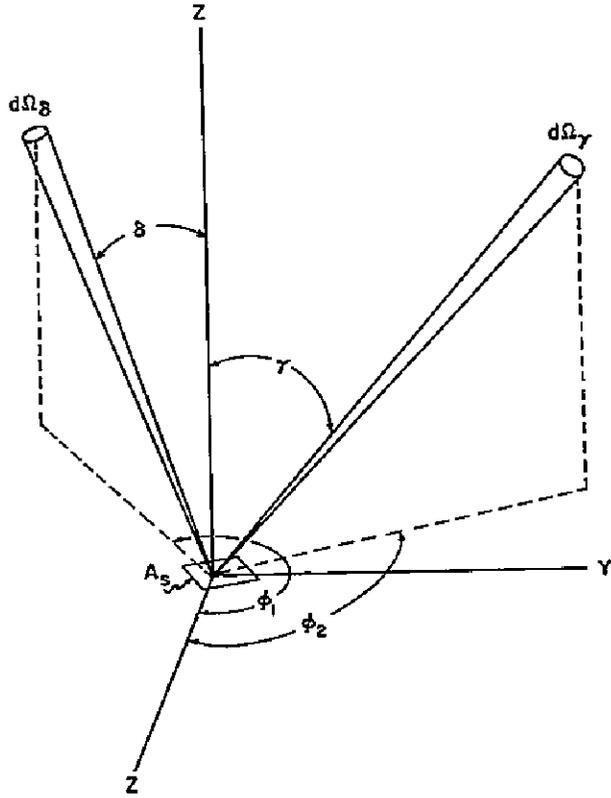


Fig. 3. Geometry of incident and reflected radiation used in bi-directional reflectance discussion.

expression for this relation (Nicodemus (1965)), is:

$$N'_S(\phi_2, \gamma) = \int \rho' N'_r(\phi_1, \delta) \cos \delta d\Omega_\delta \quad (12)$$

where the effective incoming radiance $N'_r(\phi_1, \delta)$ as well as $N'_S(\phi_2, \gamma)$ depends on an azimuth angle, ϕ , in the plane of A_S ; and $d\Omega_\delta$ is an elemental solid angle subtended by the radiation source at A_S . The integration is taken over all such solid angles.

For the special case when the incident radiation is only direct solar energy (no "sky" radiation) we may assume that ρ' and $N'_r(\phi_1, \delta)$ are constant over Ω_δ , which is the solid angle subtended by the sun. Now δ is the zenith angle of the sun and the subscript r is replaced by \odot for the sun. Then (12) becomes

$$N'_S(\phi_2, \gamma) = \rho' N'_\odot(\phi_1, \delta) \cos \delta \Omega_\odot \quad (13)$$

and since $H_{\odot}' = N_{\odot}'(\phi_1, \delta) \Omega_{\odot}$, we have

$$N_S'(\phi_2, \gamma) = \rho' \cos \delta H_{\odot}' \quad (14)$$

or

$$\rho' = \frac{N_S'(\gamma, \phi_2)}{H_{\odot}' \cos \delta} \quad (15)$$

and from Eq. (11)

$$\rho' = \frac{P'}{A_C \omega_C H_{\odot}' \cos \delta} \quad (16)$$

Note also that the effective directional reflectance, $r' = r'(\phi_1, \delta)$ is defined as the ratio of radiation incident on A_S from the direction δ, ϕ_1 , to that reflected (or scattered) in all directions and is related to ρ' by

$$r' = \int_0^{2\pi} \int_0^{\pi/2} \rho'(\delta, \phi_1, \gamma, \phi_2) \cos \gamma \sin \gamma \, d\gamma d\phi_2 \quad (17)$$

and only for a Lambert reflector, where ρ' is constant over all directions, it is possible to integrate (17), so

$$r' = \pi \rho' \quad (18)$$

and Eq. (15) becomes

$$r' = \frac{\pi H_S'(\phi_2, \gamma)}{H_{\odot}' \cos \delta} \quad (19)$$

Since many natural surfaces and clouds are known to be non-Lambert reflectors, measurements of ρ' from the ATS data can only be used to compute effective "albedo" values if a bi-directional reflectance pattern is known or assumed (see Bartman (1967)). In addition, some knowledge of the spectral reflectance properties of the regions viewed is required before the measurements in the 450 - 650 nonometer region can be used to infer the "total" short-wave albedo.

5. Prelaunch Calibration Tests

In order to relate P' to camera output, a reliable calibration must be made. This has been carried out by K. Hanson and V. Suomi. However, it is possible to use the prelaunch calibration tests as a basis for understanding the requirements for determining earth radiance values, and as an estimate of the relationship between P' and camera output.

The primary result of the prelaunch calibration tests was the determination of the required gain of the photomultiplier and the video amplifier such that the brightest clouds would not saturate the system while allowing a good dynamic range. A second important result was the determination of the slope of the input-output curve of the camera, and establishing the fact that below the saturation level this relationship is nearly linear (Fig. 4). To accomplish this, two kinds of measurements were made. In one the camera and a spot photometer with a spectral response close to that of the camera (Fig. 2) viewed a light box with a variable quartz iodine source. Plots of the photometer output against the camera output on a log-log scale showed that below the saturation level the response of the camera was linear. Another set of measurements was made outdoors with the camera and the photometer viewing a piece of Kodak white paper illuminated by sunlight at normal incidence at various times of day. (Fig. 5.) From a knowledge of the solar zenith angle at the time of each measurement, the number of optical air masses, (m), was calculated. Plots of camera output v. s. m . extrapolated to the top of the atmosphere give the photomultiplier output when viewing nonattenuated sunlight incident normal to a highly reflecting surface. This was assumed to be near the maximum input to the camera from clouds. Combined with the results of the previous test, this information was used to set the photomultiplier supply voltage and the video amplification such that this maximum value occurred at the top of the linear portion of the curve.

6. Relationship Between Camera Output and P'

The results of the prelaunch (solar) tests can also be used to illustrate a method for determining the relationship between P' and camera output if the solar spectral distribution at the time of the tests is assumed to be known. Since any linear relationship when plotted on a log-log plot has a slope of one, only one point on the slope need be calculated to establish the P' vs. camera output relationship.

The following calculation of P' is based on the spectral distribution of solar radiation for $m = 0$ from the Handbook of Geophysics and Space Environments (1965).

The geometry for the test set-up was as shown in Fig. 5. As can be seen, the source in this case is a piece of Kodak white paper (average total directional reflectance, $R = 0.88$), at a distance of only 49 inches from the camera. Considering Eq. (11) from Sec. 2, it can be seen that P' , however, is independent of the distance from the source. Thus, the radiance of the paper can be substituted into this equation to get the effective power.

The tests were made with the paper kept nearly perpendicular to the sun's rays and the camera axis nearly perpendicular to the paper, the deviation from the normal being just enough so that the camera did not cast a shadow on the paper.

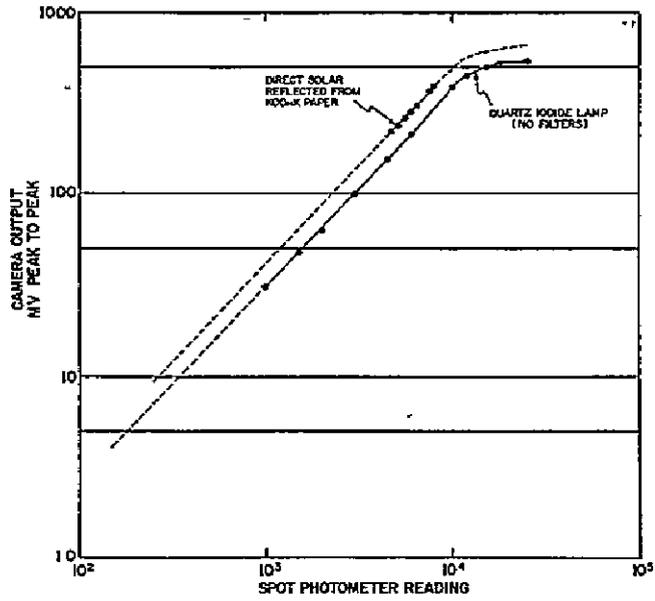


Fig. 4. Spot photometer output vs. camera output for two sources. Data taken during prelaunch tests.

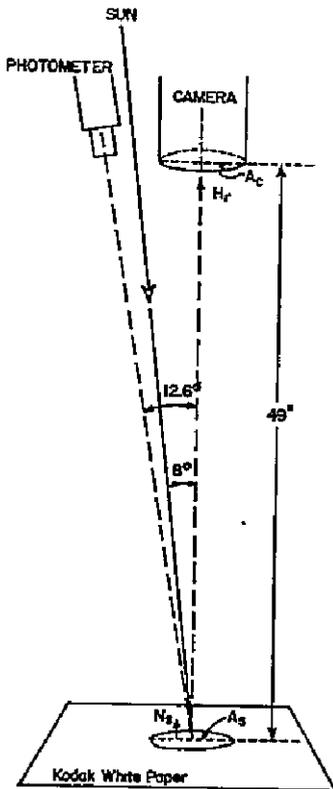


Fig. 5. Geometry of prelaunch solar test.

The effective spectral radiance of the paper is then given by:

$$N'_S = \int_{\lambda_1}^{\lambda_2} \frac{H_{\lambda\odot}}{\pi} R_\lambda r d\lambda \quad (20)$$

where $H_{\lambda\odot}$ is the solar spectral irradiance on the paper, and $H_{\lambda\odot r}/\pi$ is the normal component of the radiance from a Lambert surface of reflectivity r .

Equation 20 has been evaluated numerically using the distribution for $H_{\lambda\odot}$ at $m = 0$ and the spectral response of the camera R_λ shown in Fig. 2.* The result of this calculation is

$$P' = 4.55 \times 10^{-9} \text{ [watts]} \quad (21)$$

From Fig. 6, the camera output for this input was 700 millivolts (mv). This, then, established the point which was used to plot the curve shown in Fig. 7. As shown in Fig. 4, this value would actually saturate the camera, and so is shown on the curve on an extension of the linear portion.

It should be mentioned that this curve is only an estimate based on pre-launch calibration tests which were intended to assist in making necessary adjustments to the camera gain level setting in order that the camera would respond favorably to viewed conditions on earth.

7. Field of View Considerations

The real field of view does not have a uniform response. Its probable response was determined during the prelaunch calibration tests (Thomsen, et al.). The angular field of view was estimated to be 0.1 m.r. at the half-power points, tailing off to zero at 0.35 mr. From the response curve in the above reference, it can be seen that about 26 percent of the detected input is from beyond the 0.1 mr field.

In practice, the way to deal with this situation is influenced by the fact that both the ground calibration and the inflight calibration were done with sources that completely fill the field of view. Thus, in order to use these calibration curves the data used must represent a source which also entirely fills the field of view.

*See paper by Thomsen, Parent and Suomi for a tabulation of R_λ and a more detailed plot.

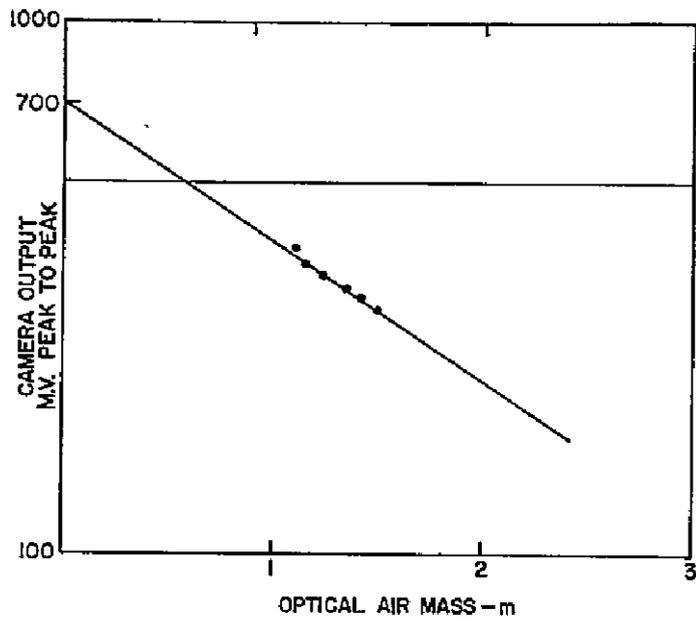


Fig. 6. Camera output viewing sunlight vs. number of optical air masses.

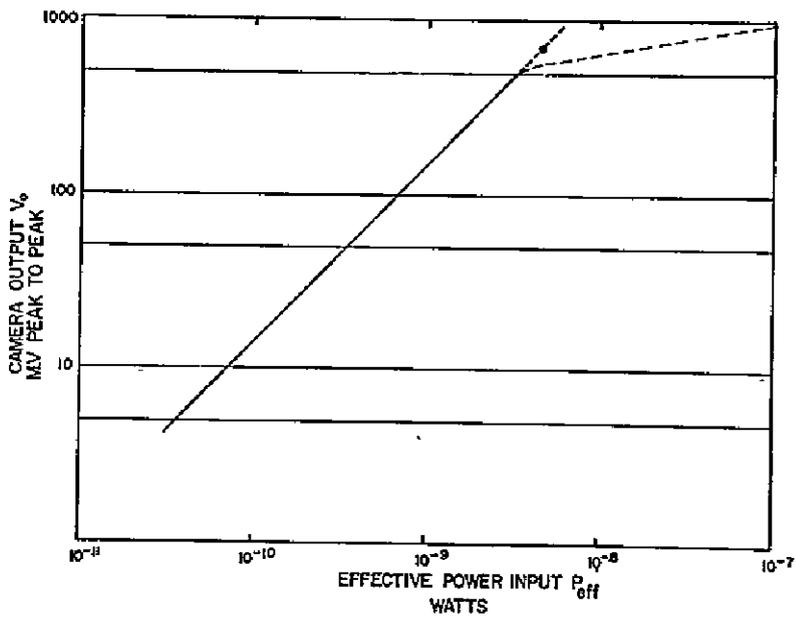


Fig. 7. Camera output vs. effective power input.

This is especially important to note for those who wish to measure the radiance from small clouds. The following examples, shown in Figs. 8 and 9, are intended to demonstrate ambiguities and other difficulties which arise when one attempts to obtain a measurement of a source smaller than the field of view from the digital ATS data.

In the first set (Fig. 8), a uniform field of view is assumed. The signals from clouds of equal radiances larger than or equal to the field of view have equal amplitude. However, a cloud of the same radiance smaller than the field of view does not produce the same amplitude, because the signal is integrated over the whole field. The signal from a cloud which does not have as large a radiance, but fills more of the field could have the same amplitude as the smaller brighter cloud, but a different duration. In addition, the duration of the signal produced by clouds of the same size but smaller than the field of view is different depending on their location within the field of view in the direction perpendicular to the scan direction.

In the second set, a field of view more like the real case is assumed. In this case the same difficulties arise as for the step function field of view, with added ambiguities. Here a cloud seen by the "outer edge" of the detector which is less sensitive could have the same amplitude as a less bright cloud seen by the center of the detector.

From these examples, it is clear that the only time one can make unambiguous measurements of the radiation from a cloud is when it fills the whole field of view (i. e., beyond 0.35 mr). To get an idea of how large a cloud has to be to get a measurement of its effective radiance by the camera, consider the plot shown in Fig. 10. Here the major and minor axes and the area of the ellipse formed by the intersection of a 0.1 mr cone, vertex at the satellite, and a plane tangent to the earth at the intersection of the earth and the axis of the cone is plotted against the nadir angle of the camera.

Clouds having these areas at these distances would fill the field of view defined by the half power points, i. e., the amplitudes of the signals from them would be approximately 70 percent of the amplitude for clouds of the same effective radiance filling a 0.35 mr. field of view.

If one were using the digital data, one would have to average the number of digital samples in the field of view. The signal is digitized linearly with respect to the scan angle, and there are 8192 samples within a 20° scan. This means a sampling rate of about 2.35 samples per 0.1 milliradian, or about 8.23 samples per .35 milliradians. Thus, measurements from areas smaller than 10 samples wide (8 n. mi. sq. at the subpoint and 15 n. mi. sq. near the limb) must be interpreted with care when using the calibration curves.

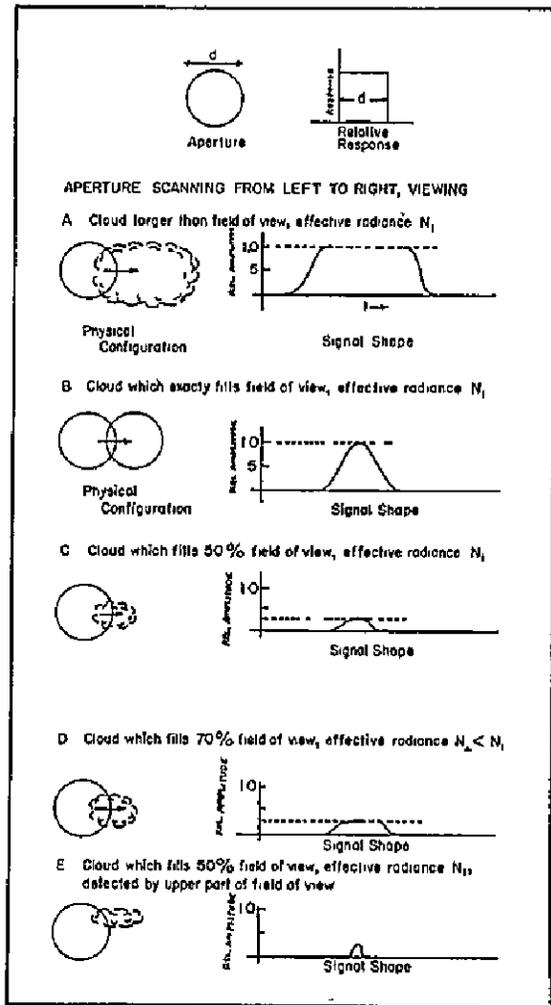


Fig. 8. Schematic representation of signals from various sources as seen by an aperture having a uniform field of view.

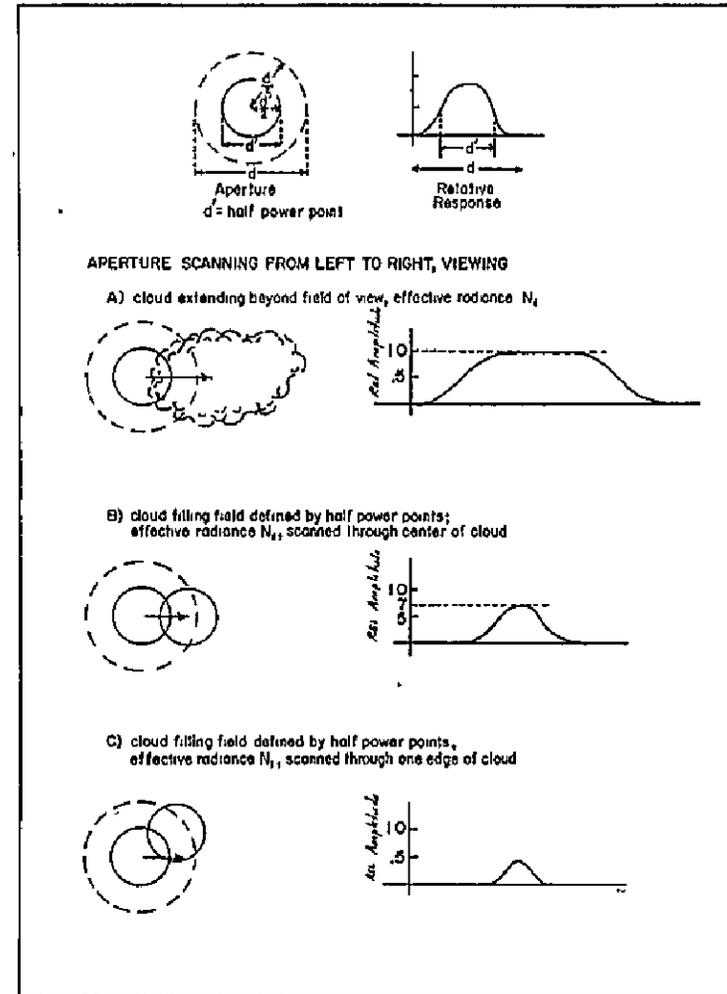


Fig. 9. Schematic representation of signals from various sources as seen by an aperture having a nonuniform field of view.

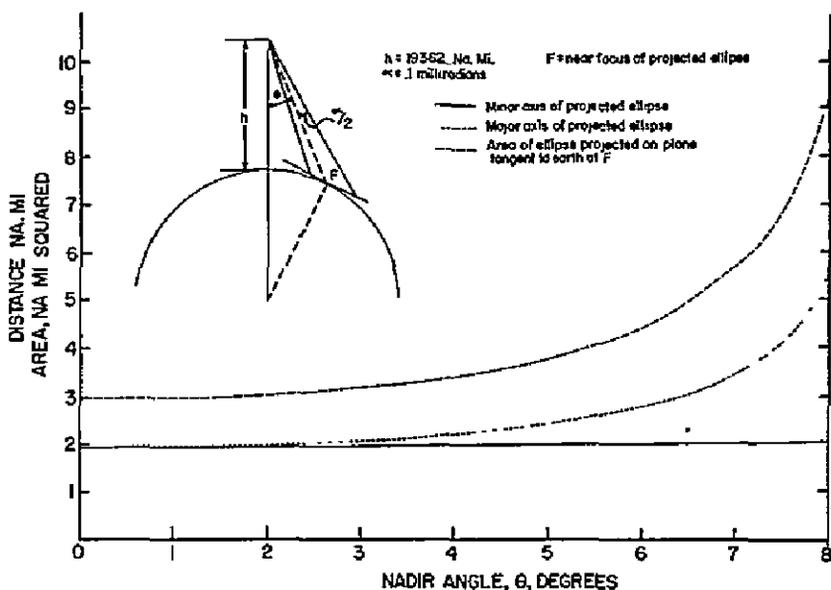


Fig. 10. Camera nadir angle vs. area, major axis and minor axis of projected ellipse.

8. Using the Digital Data

When using the digital data with the calibration curve in Fig. 4, the digital count (or average digital count) must be converted to millivolts out of the camera.

The video amplifier on board the satellite has two gain modes which can be commanded from the ground: mode 2, or normal mode, and mode 1, or high gain mode, which is 10 db* (3.161 times) higher than mode 2. In addition to this, there is the capability of increasing the gain on the ground up to 12 db higher (3.98 times) in steps of 2 db (1.26 times).** Table 1 lists these various gain steps and their corresponding amplification factors.

*The notation db means decibels. The number of decibels is equal to $20 \log V_2/V_1$, where V_2/V_1 is the ratio of the voltage outputs.

**This gain information is recorded on the digital tapes as part of the documentation code.

Table 1

Spacecraft Gain	Ground Station Gain	Amplification Factor Relative to Lowest Setting
MODE 2 — 0 db	0 db	1.00
"	2 db	1.26
"	4 db	1.58
"	6 db	1.99
"	8 db	2.51
"	10 db	3.16
"	12 db	3.98
MODE 1 — 10 db	0 db	3.16
"	2 db	3.98
"	4 db	5.01
"	6 db	6.31
"	8 db	7.94
"	10 db	10.00
"	12 db	12.59

As a result of the prelaunch test, the gains were set so that the maximum anticipated output of the camera video amplifier, when viewing bright clouds, would be approximately 500 millivolts in normal gain mode, zero ground gain. The telemetry link from the satellite through the ground station analog to digital (A/D) converter, is set so that the maximum digital number, $D = 255$, occurs for a 500 millivolt signal from the satellite video amplifier when in mode 2, zero ground gain.

Thus if the camera is in the high gain mode and/or the ground gain is other than zero, the voltage represented by the digital value must be adjusted to give the equivalent reading for mode 2, zero ground gain, in order to use the calibration curve done in this gain.

For example, assume the camera is in high gain mode, the ground gain is 2 db, and the digital number for a certain camera reading is 100. Then

$$\frac{\text{millivolts}}{\text{count}} = \frac{500}{255} = 1.96$$

$$1.96 \times 100 = 196 \text{ m. v.}$$

The high gain mode is 3.161 times higher than the normal mode, so

$$\frac{196}{3.16} = 62 \text{ m. v.}$$

is the corresponding reading for normal mode, ground gain 2 db. To convert to ground gain zero, this must be divided by 1.26 (2 db), which gives

$$\frac{62}{1.26} = 49.2 \text{ m. v.}$$

as the millivolt output of the camera referred to mode 2, zero ground gain. Hanson and Suomi have referenced their inflight calibration to mode 2, ground gain 0.

9. Summary

Measurements of $N_S'(\gamma)$ can be derived from the ATS-I observations, and the relationship between these reflected radiance values and the camera voltage output or digital counts has been shown. Even with the camera inflight calibration, the digitized ATS data must always be considered in terms of the following:

- a) the camera measures only the effective radiance, $N_S'(\gamma)$ between 450 and 650 $m\mu$ reflected and scattered through the "top of the atmosphere" in a given direction.
- b) the camera has a non-uniform response across its nominal field of view and thus measurements over areas of nonhomogeneous radiance must be interpreted with care.
- c) the relation between the digital count values and the camera output voltage depends on two gain settings and must be considered before calibration curves are used to obtain absolute values of N_S' .

In addition, the ATS-I data can be used to measure the effective bi-directional reflectance of the region viewed by the camera. However, the conversion of this measurement into an albedo value, even for the instant of observation, requires supporting information on the bi-directional and spectral reflectance pattern of the area in view.

REFERENCES

1. Bartman, F. L., 1967: The reflectance and scattering of solar radiation by the earth, Technical Report 05863-11-T, The University of Michigan, College of Engineering, 257 pp.
2. Handbook of Geophysics and Space Environments, Valley, S. L. (ed.), 1965: McGraw-Hill, N. Y., 16-4, 16-6.
3. Hanson, K. J. and V. E. Suomi, 1968: In-flight calibration of the ATS-I camera. ATS-I Volume, University of Wisconsin Press.

4. Nicodemus, F. E., 1965: Directional reflectance and emissivity of an opaque surface, Appl. Optics, 4, 767-773.
5. Suomi, V. and R. Parent, 1968: Calibration procedure for ATS-I spin-scan camera. ATS-I Volume, University of Wisconsin Press.
6. Thomsen, R., R. Parent and V. Suomi, 1968: The ATS-I camera. ATS-I Volume, University of Wisconsin Press.

A STUDY OF CLOUD DISTRIBUTIONS USING REFLECTED RADIANCE
MEASUREMENTS FROM THE ATS SATELLITES*

Alfred J. Stamm and Thomas H. Vonder Haar
The University of Wisconsin
Madison, Wisconsin

ABSTRACT

Reflected radiance measurements from the Multicolor Spin-Scan Cloud Camera on Applications Technology Satellite III are used to determine the percentages of selected areas of the earth that are cloudfree. The areas chosen are meteorologically active and represent common cloud patterns. Use of several data unit sizes shows how the observed percent clear area decreases with decreasing spatial resolution of a simulated sensor. Methods of determining a cloud-no cloud threshold are discussed. The change of cloud cover over a period of a few hours is examined. It is found that clouds smaller than the instantaneous field of view are often not recognized as clouds and therefore tend to affect the interpretation of spacecraft camera measurements. The results of this investigation are used to suggest the optimum spatial resolution for radiometrically sounding the atmosphere from a geosynchronous satellite using an instrument described in the report.

1. Introduction

There is increasing interest today in detecting, from satellites, radiation that is reflected or emitted from the earth. In order to see the earth, a satellite must look between the clouds, but there is very little information as to the size and frequency of the cloudless areas.

One need for this information is to specify the optimum field of view of a radiometer designed to measure radiances from which the vertical temperature profile can be calculated (COSPAR Report, 1969). In order to acquire a temperature profile down to the ground, cloudless areas are preferred, but the most

*Research sponsored by NASA Contract NAS 5-11542

interesting weather occurs where there are clouds. Since a number of readings are needed for "good statistics," especially from geosynchronous orbit, the problem of how close one can sound to baroclinic meteorological regions translates to the problem of what percentage of the area near various cloud patterns is cloudless. Also needed is information on the size of the cloudless areas, or rather what instrument spatial resolution is needed to see between the clouds. This is necessary because a tradeoff must be made between the spatial resolution and the sensitivity of the instrument. More will be said about sounding in a later section.

Scientists in the Earth Resources Technology Satellite (ERTS) program are interested in looking at certain areas of the earth with various sun angles and at various times of the year. Salomonson (1969) gives an extensive list of references on cloud statistics. Before a viewing program can be outlined, information is needed on the clouds covering these areas at the times of interest, especially areas that are cloudy most of the time. Such questions as: "Can the information be obtained by looking between the clouds, or are cloudless days needed?", and "what is the likelihood of a cloud cover at various times of the day and year?", will need to be answered.

A third possible need for cloud statistics data could be in determining the types of clouds by correlating the cloud radiance and variance in radiance with the amount of cloudless area. This approach may be called a quantitative nephanalysis. The present study is an initial attempt to derive information on cloud coverage specifically needed for the vertical temperature sounding problem. Portions of the work presented apply equally well to other applications of radiance measurements from a geosynchronous platform.

2. Description of Data

Among the best cloud cover data sources available today are the Applications Technology Satellite (ATS) spin-scan cloud camera pictures from ATS-I and ATS-III. In these experiments the "camera" consists of a photomultiplier behind the optics of a small telescope. The pictures are made as the instantaneous field of view (IFOV) of the camera telescope (0.1 mrad) scans across the earth while the satellite in geosynchronous orbit is spinning on an axis parallel to the earth's axis (Thomsen, Parent, and Suomi, 1969). After each scan, while the camera is pointed away from the earth, the camera is stepped down by an amount slightly greater than the angular field of view so that a new scan is made parallel to and just south of the previous scan. On ATS-III, 2400 scans north to south comprise a complete picture of the earth. At the nominal 100 rpm spin rate this means a picture can be made in slightly less than one-half hour. The analog camera output of the 20° segment which comprises a scan across the earth is digitized into 8192 samples of 2.35 samples per IFOV. Each sample is made up of 8 bits giving a dynamic range of 256

radiance levels. Vonder Haar (1969) discusses the relationship between the digitized counts (relative radiance) and the reflected solar radiance measured by the photomultiplier.

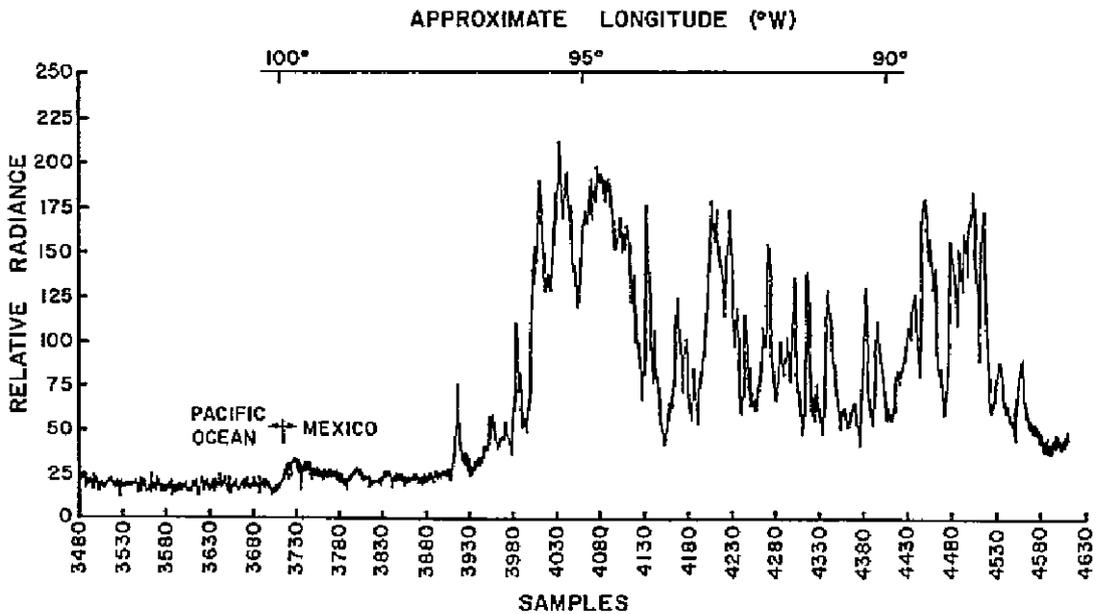


Fig. 1. Radiance level plot for part of a single scan line. January 20, 1968 (191131Z). Green channel. Latitude about 17°N.

Figure 1 shows a radiance level plot for part of a single scan line. ATS-III pictures have three colors—blue, green, and red—but only the green channel (0.48 - 0.58 μm) is used in the work reported in this article. This figure shows very well both the high dynamic range and high spatial resolution of the ATS data.

Since there are more than 2 samples per field of view, each sample does not necessarily represent new information (i. e., the scan samples overlap). For this reason and also in order to work with a more square basic unit size, the values of 3 adjacent samples are averaged. Using the scan step size rather than the field of view, the size of the basic unit is roughly 4.6 km north to south at nadir and slightly less east to west. These basic units are then combined so that 1, 4, 9, 16, 25, or 36 go into a square data unit which is the area corresponding to the simulated spatial resolution or instantaneous field of view of a future instrument. A grid or field is the square area over which the computations are made and is 60 basic units across for the former and 100 basic units across for the latter. This corresponds to 276 km and 460 km at nadir respectively. The former size was chosen because it is divisible by 1, 2, 3, 4, 5, and 6, and the latter size was used for parallel work

being done on the same ATS picture (Hanson, 1969). Note that these areas are 48% and 132% respectively of the area of a basic Global Atmospheric Research Program (GARP) grid mesh (GARP Report, 1969). A matrix is that combination of grids or fields covering the particular cloud pattern of interest.

3. Methods of Simulating the IFOV

In terms of the ATS data, the problem is stated as follows: What percentage of the different size data units within a given grid or field are cloudfree? Two methods have been used to treat this problem.

The first method, called Method A, approximates an instrument which uses its entire IFOV to compare with a cloud-no cloud threshold (T_h). This threshold, which is given in terms of a radiance level, will be discussed further in the next section. Method A uses the square data units and takes the average radiance of all the samples within the data unit to compare with the threshold. That is, each data unit is clear that satisfies the equation:

$$T_h - \frac{\sum_{i=1}^n S_i}{n} > 0$$

where S_i is the radiance value of the i^{th} sample within the data unit and $n = 3, 12, 27, 48, 75, \text{ or } 108$ depending on which data unit size is used. The observed percent clear area is then the percentage of the data units within a grid or field that satisfy the above equation (i. e., are entirely clear). This percentage will change with spatial resolution simulated by changing the data unit size. It should be noted here that the percent clear area observed at a given resolution is not the actual percent clear area (i. e., infinite resolution).

Figure 2 shows that when using Method A, the observed present clear area does not decrease significantly with deteriorating spatial resolution (i. e., increasing data unit size). In some cases it even increases. This is misleading because as the IFOV increases, small clouds are more likely to be found within the IFOV. In other words, viewing between some of the clouds may no longer be possible, and the observed percent clear area should decrease.

The reason for this difficulty is as follows. When averaging over a large data unit, a small cloud may be obscured since a small fairly dark cloud can average with dark land around the cloud to equal a radiance level comparable to bright land. (In ATS pictures, land and ocean are generally darker than clouds.) As a result, clouds smaller than the IFOV may not be detected.

This difficulty is significant for the following reason. Based on experience with the relatively low spatial resolution temperature sounders on NIMBUS-III, Smith* estimates that ten percent cloudiness within the IFOV is enough to

*William L. Smith, NESG, ESSA, private communication.

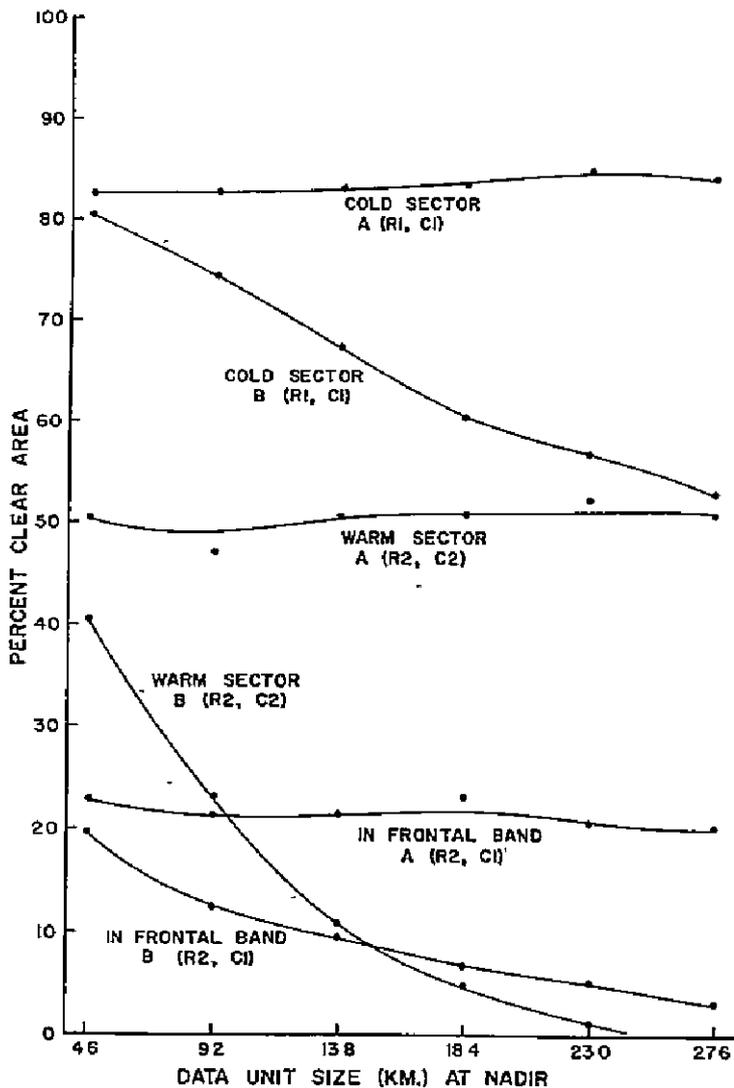


Fig. 2. Percent clear area for various data unit sizes of an extratropical front over Northwestern United States. April 23, 1968 (171410Z) Matrix 3. A and B refer to method (see text) and row and column locate grid within matrix.

make a temperature profile below the cloud level inaccurate. Smith (1968) also shows that the effect of such partial cloudiness can be overcome only with the aid of special cloud statistics or additional information about the atmospheric temperature profile.

To alleviate the difficulty of averaging over large data units (Method A), a second Method, B, was designed. It labels data units as cloudy even if

only one sample is over the threshold. Method B, therefore, simply changes every radiance level over the threshold to one million and then proceeds as in Method A. The problem of calling small dark clouds "clear" is thus almost eliminated because the averaging is done over a much smaller area (i. e., an ATS sample). This method approximates an instrument that uses another channel with an IFOV smaller than the sounding channels (e. g., a cloud mapping channel used to determine cloud velocities) to determine whether or not clouds are actually present. In other words, a data unit is now observed as clear if it satisfies the following equation:

$$\text{Th} - \frac{\sum_{i=1}^n r_i}{n} > 0$$

where $r_i = S_i$ if $\text{Th} \geq S_i$

$r_i = 1,000,000$ if $\text{Th} < S_i$.

(It should be remembered that S_i ranges from 0 to 255.)

In comparing Method A with Method B (Fig. 2), note that not only is the decrease in percentage clear area with resolution greater for Method B than for Method A, but also that the percentage clear area at 4.6 km resolution is less for Method B. The latter occurs because Method A averages over three samples while Method B compares the value of each sample with the threshold. This gives some insight as to how an actual instrument (including, of course, the ATS sensor*) may erroneously flag an area as cloudfree, when clouds smaller than the IFOV are present. Since most instruments approximate Method A (i. e., they average the radiances within their field of view), an instrument to be used for radiometrically sounding the atmosphere should be designed so that the cloud recognition is made at a higher resolution than the sounding.

Figure 3 shows a map of the clear areas at three spatial resolutions using Method B. It depicts a case of cumulus and stratus clouds in the subtropics. If this region had been treated with Method A, not all of the 9.2 km resolution cloudy area would have been within the 18.4 km resolution cloudy area nor the 4.6 km resolution area within the 9.2 km resolution area. The remainder of this study uses Method B.

4. Threshold Determination

As mentioned before, each element has 256 possible radiance levels. Clouds are almost always brighter than either land or water but the question is:

*W. Shenk, GSFC, NASA, is currently comparing the ATS radiance measurements of cloud fields with even higher resolution data obtained from simultaneous Apollo photographs.

"How much brighter?" Cirrus and other type clouds may be quite dark (poor reflectors). This makes it difficult to determine a cloud-no cloud threshold.

A combination of two methods is used in this paper to select thresholds. First a plot of relative radiance versus sample number is made for a given line. By locating this line on the corresponding picture, the samples which represent cloudless land and ocean values can be determined. The upper values of clear land and ocean are then extracted from the line plot. Figure 1 shows a typical line plot of ATS-III. It should be noted that in the green channel (as stated before, the green channel is the only one used in this experiment), the ocean is darker than the land. This implies that the difference in radiance levels between ocean and clouds is greater than that between land and clouds. For this reason and also because the radiance spread is less over the ocean than over the land (see Fig. 1), it is easier to make a threshold determination over the ocean.

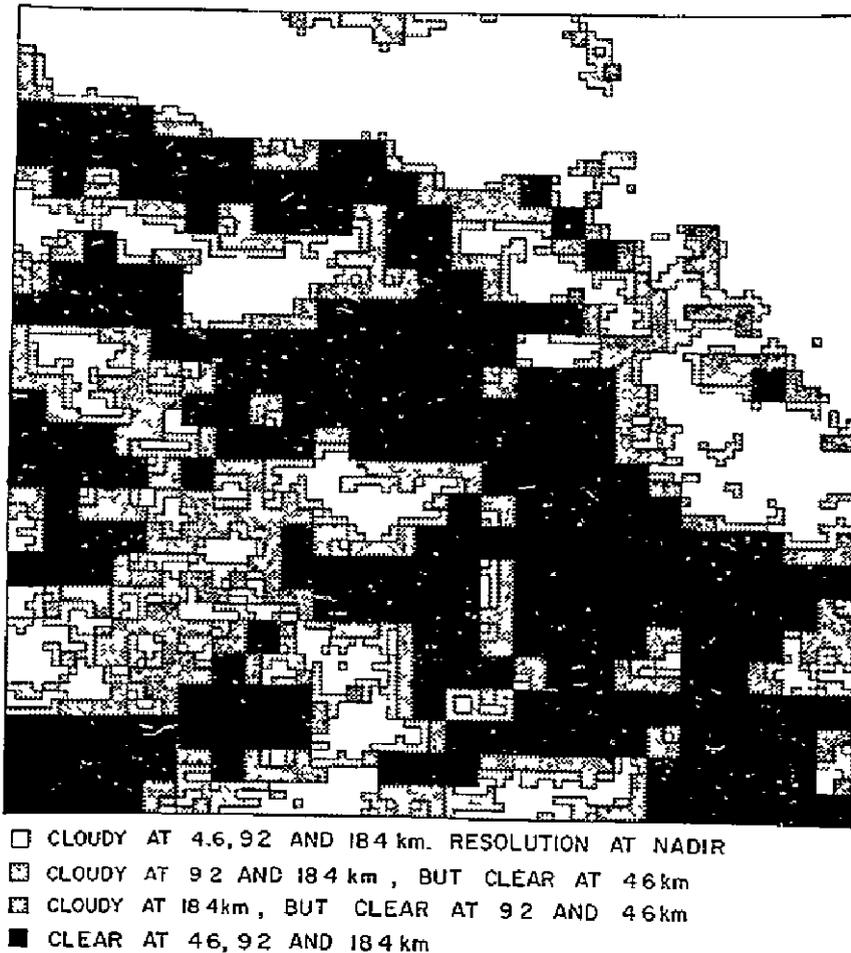


Fig. 3. Map of cloud cover for a high pressure zone cumulus and stratus in the subtropics. January 20, 1968 (191131Z) Field 85.

The second method of determining the cloud-no cloud threshold is to plot the frequency at which a given radiance level occurs within a grid. If there is a reasonable amount of clear area within the grid, a large peak of roughly Gaussian shape will represent this clear area provided the clear area is homogeneous as is generally true over oceans. The point at which the curve starts to rise again or a few radiance levels to the cloudy side of this clear area peak is taken as the point where cloud interference is becoming important. This point is chosen as the threshold. Figure 4 shows a radiance level frequency plot for a cloudy area over the ocean. Land areas lack the homogeneity of oceans so this method of threshold determination is more difficult. Figure 5 shows a radiance level frequency plot for the average of eight grids over land.

In order to check the significance of an incorrect threshold, these eight grids were run with five different thresholds. The additional thresholds are indicated by small arrows in Fig. 5. Figure 6 shows the average clear area of all eight grids in the matrix versus resolution for each of these thresholds. It is seen that the percentage clear area changes greatly with a change in threshold but that the shape of the curve remains about the same. Since it is difficult to obtain the threshold to within several radiance levels, an error might be expected in the values of percent clear area given in this paper; however, the change in percentage open area with resolution is more significant and also more accurate, and will therefore be emphasized in this paper.

In summary, cloud-no cloud thresholds can be determined from several methods. Two of the most basic have been used in this work. Particularly over

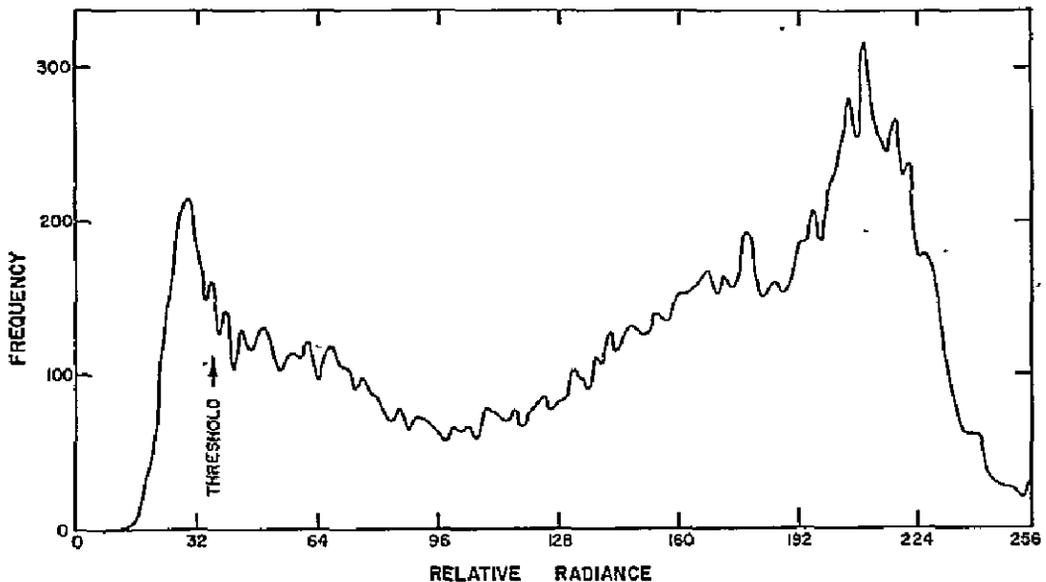


Fig. 4. Frequency plot of radiance levels over water. Raw data corrected according to Hanson (1969). January 20, 1968 (191131Z) Field 94. Thirty-six was taken as threshold.

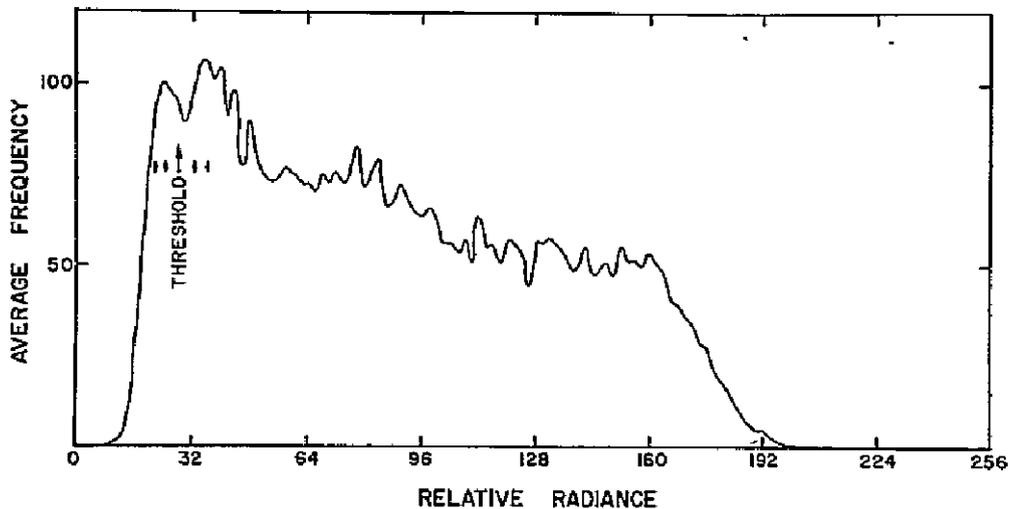


Fig. 5. Average frequency plot of radiance levels for eight grids over land. Raw data corrected according to Hanson (1969). April 23, 1968 (171410Z) Matrix 1. Twenty-eight was taken as threshold. Small arrows indicate additional thresholds used in Fig. 6.

land areas the frequency distribution technique is difficult; thus more reliance was placed on the line plot displays in selecting a threshold over land areas. In all cases, a conservative estimate of the thresholds was attempted (i. e., when in doubt, the threshold was chosen to overestimate the cloudy regions).

5. Results

ATS measurements from April 23, 1968 (171410Z) and January 20, 1968 (191131Z) were chosen for this study. The grids for which the results are listed in this paper are in Matrices 3 and 5, and part of Matrix 1 of Fig. 7, and fields 37-40, 79-81, 83-85, and 87-95 of Fig. 8. Table 1 gives the percentage clear area using Method B for each of the grids mentioned above. These six matrices represent various meteorological conditions including extratropical fronts, cumulus and cumulonimbus over the Amazon, cumulus and stratus in the subtropical high pressure zone over the Pacific, and the inter-tropical convergence zone.

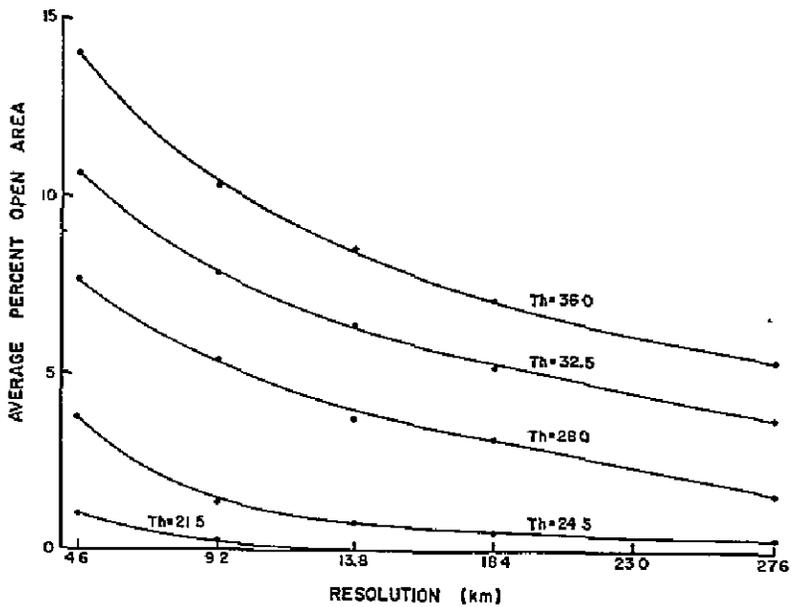


Fig. 6. Comparison of 5 thresholds for the same grids described in Fig. 5. The percent clear area is given at different spatial resolutions for the average of these grids.

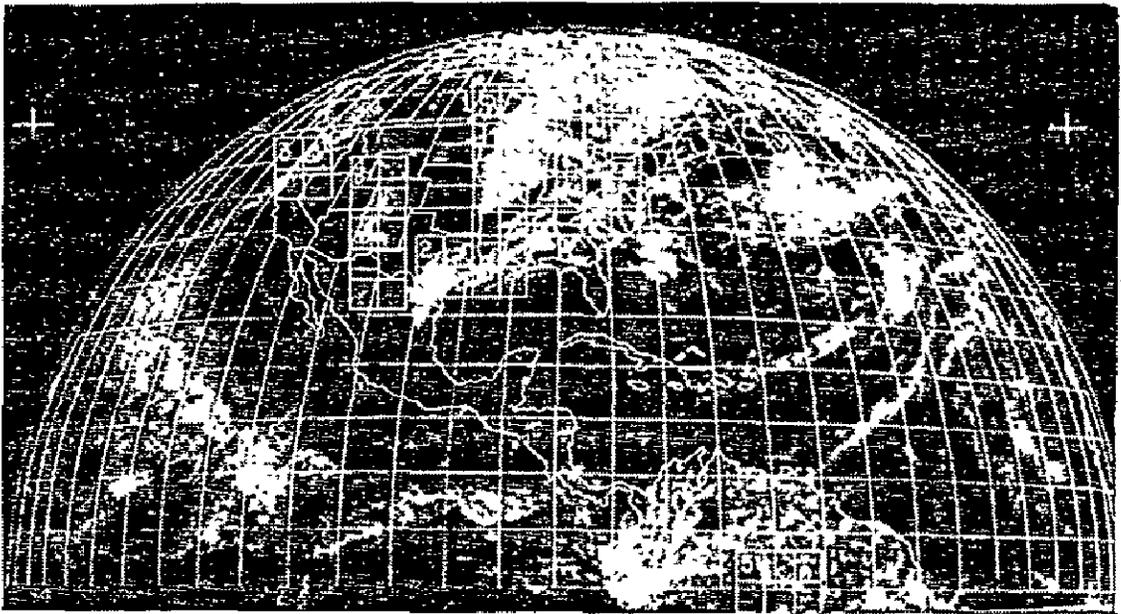


Fig. 7. April 23, 1968 (17I410Z). Investigated areas include Matrix 1, rows 4 and 5, columns 2-5; Matrix 3 and Matrix 5.

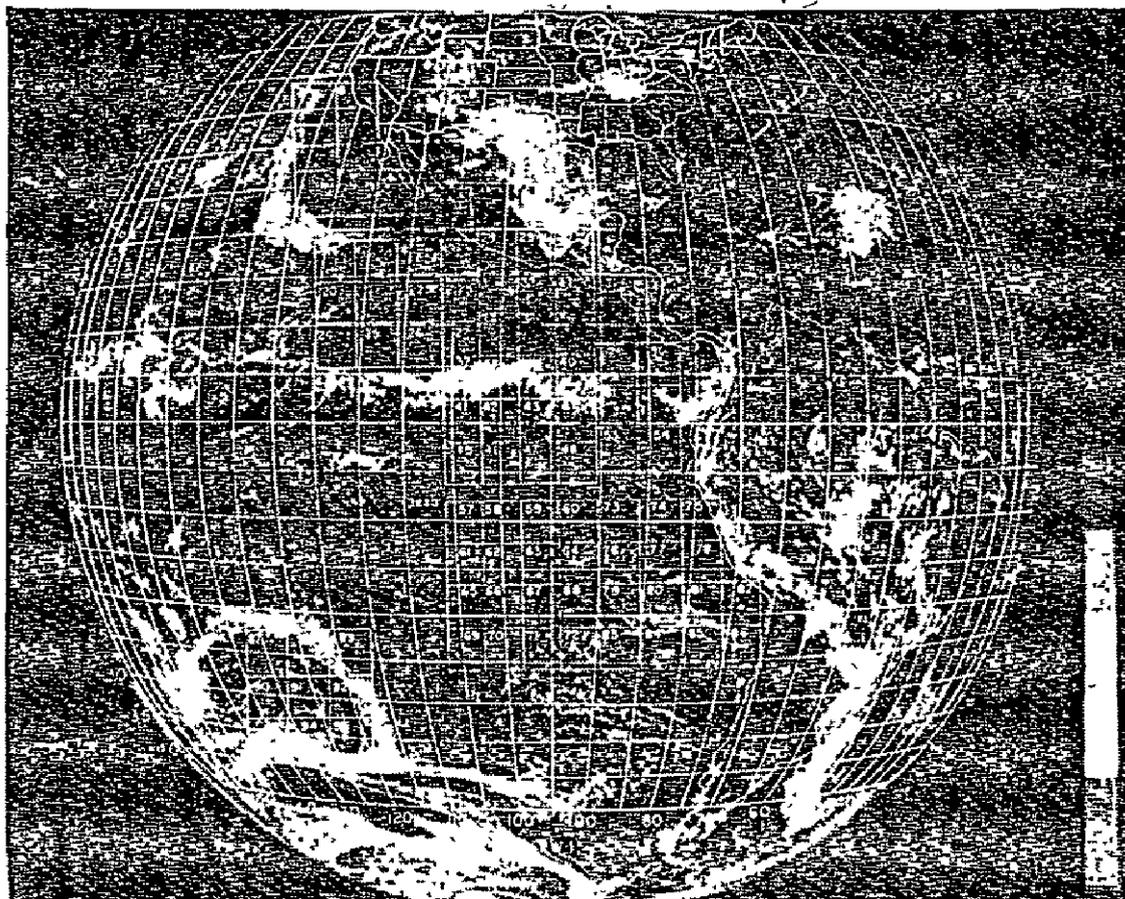


Fig. 8. January 20, 1968 (191131Z). Investigated areas include Fields 37-40, 79-81, 83-85, and 87-95.

TABLE 1

Percentage Clear Area Obtained by Method B

April 23, 1968 (171410Z) Matrix 1
Extratropical Front over Midwestern U. S.

Resolution at Nadir (km)	4.6	9.2	13.8	18.4	23.0	27.6
Row 4 cl 2	11.36	9.00	7.25	5.78	4.17	4.00
Row 4 cl 3	16.17	10.00	5.50	4.00	2.78	0
Row 4 cl 4	0	0	0	0	0	0
Row 4 cl 5	8.00	4.22	1.75	1.33	0	0
Row 5 cl 2	24.00	19.33	15.25	13.78	8.33	9.00
Row 5 cl 3	0.53	0.22	0	0	0	0
Row 5 cl 4	1.14	0.11	0	0	0	0
Row 5 cl 5	0	0	0	0	0	0

(Table 1 continued)

(Table 1 continued)

April 23, 1968 (171410Z) Matrix 3
 Extratropical Front over Northwestern U. S. and Ocean

Resolution at Nadir (km)	4.6	9.2	13.8	18.4	23.0	27.6
Row 1 cl 1	80.33	74.22	67.75	60.89	56.94	53.00
Row 1 cl 2	8.92	6.33	4.75	3.11	2.08	1.00
Row 2 cl 1	18.89	12.67	9.50	6.67	5.56	3.00
Row 2 cl 2	40.53	23.11	10.25	4.89	1.39	0

April 23, 1968 (171410Z) Matrix 5
 Cumulus and Cumulonimbus over Amazon

Row 1 cl 1	9.81	2.44	0.75	0	0	0
Row 1 cl 2	9.58	2.89	0.50	0	0	0
Row 1 cl 3	12.97	3.11	0.75	0	0	0
Row 2 cl 1	41.14	15.44	6.75	2.22	1.39	1.00
Row 2 cl 2	20.11	10.89	5.75	4.44	3.47	1.00
Row 2 cl 3	12.75	3.89	0.75	0.44	0	0

Jan. 20, 1968 (191131Z) Matrix 1
 Intertropical Convergence Zone

Resolution at Nadir (km)	4.6	9.2	18.4
Field 37	7.85	5.44	3.04
Field 38	2.04	0.52	0
Field 39	4.06	1.40	0.32
Field 40	18.58	13.16	6.56

Jan. 20, 1968 (191131Z) Matrix 2
 Subtropical High Pressure Zone Cumulus and Stratus

Field 79	9.33	2.96	0.32
Field 80	6.75	4.72	2.72
Field 81	40.36	34.32	27.36
Field 83	36.70	25.52	11.84
Field 84	70.54	60.64	44.64
Field 85	65.00	54.92	39.20

(Table 1 continued)

(Table 1 continued)

Jan. 20, 1968 (191131Z)		Matrix 3	
Extratropical Front over Southwest Pacific			
Resolution at			
Nadir (km)	4.6	9.2	18.4
Field 87	61.15	50.64	38.08
Field 88	7.28	3.60	1.60
Field 89	27.47	18.40	9.76
Field 90	13.76	4.64	0.32
Field 91	24.20	13.80	6.24
Field 92	27.77	19.56	9.76
Field 93	0	0	0
Field 94	7.12	4.60	1.76
Field 95	14.60	8.72	3.52

Figures 9 and 10 compare the various grids using the clear areas derived from Method B normalized to the distribution density for each grid (defined as percent clear area at 4.6 km resolution). Use of this parameter, distribution density, provides a base to compare grids having different meteorological conditions. Figure 9 shows that although the warm sector has fewer clouds than the frontal band at 4.6 km resolution, its percent clear area decreases faster with reduced resolution. This is because the clouds in the warm sector are a more irregular mixture of different types. Figure 10, on the other hand, shows that when clouds of the same type are found in a number of grids, as is the case of high pressure zone cumulus and stratus, the decrease with resolution depends more on the distribution density. This suggests that for the same distribution density, a measure of variation in cloud type between two grids may be inferred from the rate of change of percent clear area with sensor resolution.

Figures 11 and 12 summarize our results over various types of clouds and over regions with varying percent clear area at 4.6 km resolution. Although the statistics are not large the results, as anticipated, show that both parameters are important.

In order to see how the cloud cover in the areas studied changed in several hours, pictures were selected that were 2 3/4 hours later on April 23, 1968 (195920Z) and 2 3/4 hours earlier on January 20, 1968 (162940Z). Since the selection of the threshold is very critical (see Fig. 5), it was decided to develop two selection methods that choose a threshold relative to the previous picture rather than use the threshold determination methods mentioned in the last section.

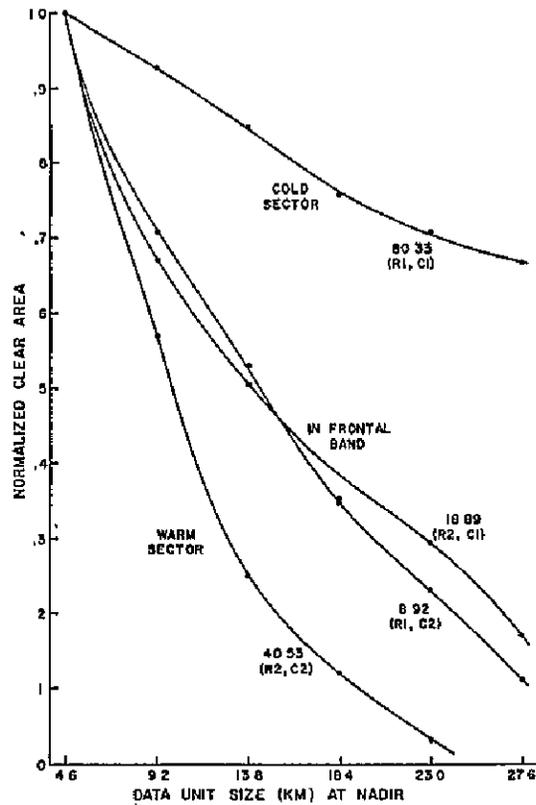


Fig. 9. Clear area versus resolution normalized to percent clear area at 4.6 km for an extratropical front over Northwestern United States. April 23, 1968 (171410Z) Matrix 3. Numbers indicate percent clear area at 4.6 km resolution and row and column locate grid within matrix.

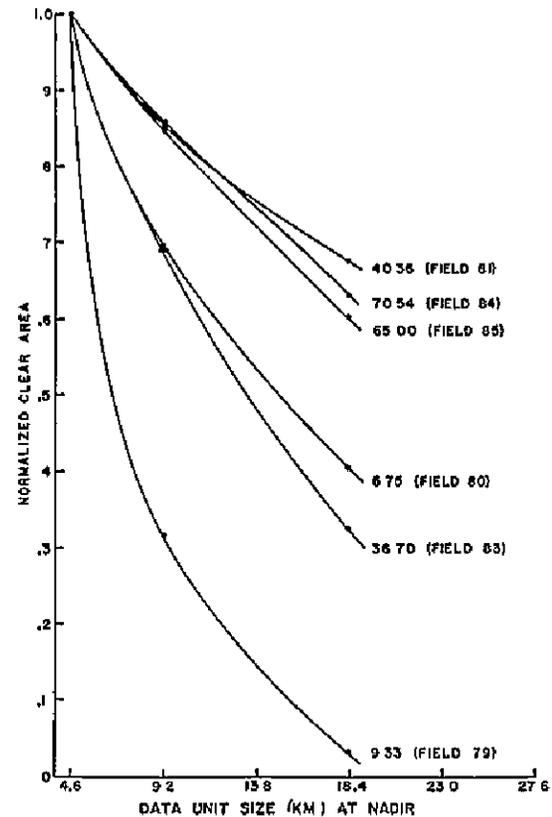


Fig. 10. Clear area versus resolution normalized to percent clear area at 4.6 km for subtropical high pressure zone cumulus and stratus. January 20, 1968 (191131Z) Fields 79-81 and 83-85. Numbers indicate percent clear area at 4.6 km resolution.

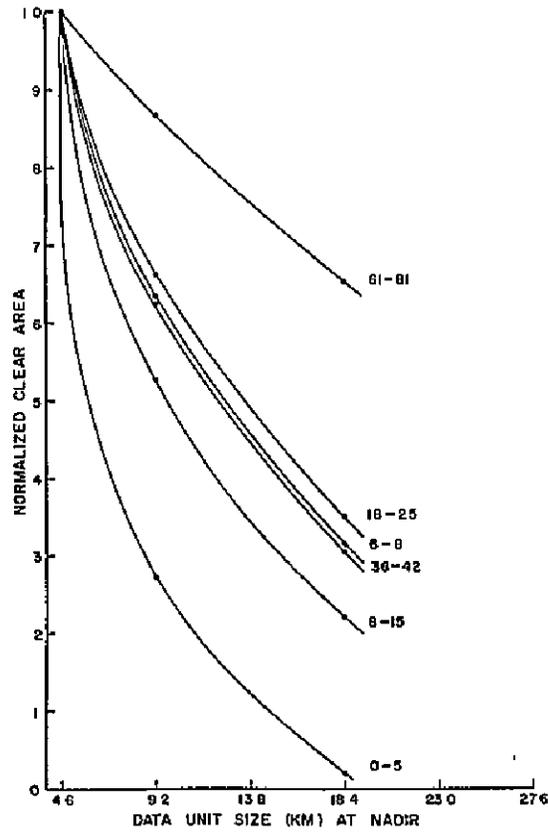


Fig. 11. Average clear area versus resolution normalized to percent clear area at 4.6 km for various degrees of cloudiness. Numbers indicate range of percent clear area at 4.6 km.

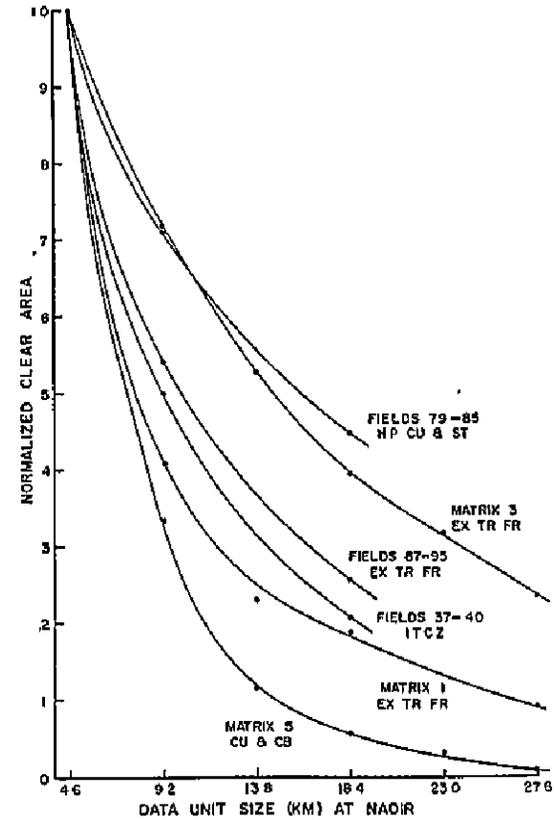


Fig. 12. Average clear area of each matrix versus resolution normalized to percent clear area at 4.6 km.

The first method uses the cosine of the sun angle at the subsatellite point as a factor for normalizing the thresholds to the pictures. In using this method, it is important that the camera and ground station gain settings are the same for the various pictures involved. Ideally each sample of each picture should be normalized using a generalized bi-directional reflectance relationship, especially when angular sun and viewing conditions vary a great deal between a pair of pictures. Such a technique is still under development.

The second method involves determining the fractional shift of the clear area peak of the radiance level frequency distribution (see Fig. 4). The peak is used because the peak is easier to find accurately than the point where cloud interference starts to dominate. (Hanson, 1969, shows that over the ocean the clear area peak can be accurately predicted by forming a Gaussian curve from the data on the left hand side of the peak—i. e., the side free from cloud interference.) Results of the two methods agree with each other to within about three radiance levels (on a scale of 0 - 255).

Figure 13 shows the change with time of the percentage of clear observations for the various fields within the intertropical convergence zone and the high pressure zone cumulus and stratus region. The total of all three types of shading must be taken to get the percent clear area of 4.6 km resolution, and similarly the total of two shadings for the 9.2 km resolution. Both of the regions studied for time variations are ones where air mass changes over a 3-hour period are negligible. Despite this fact, a wide range in both relative and absolute changes of percent clear area over the same physical area are observed.

6. An Application to the Radiometric Temperature Sounding Problem

The sounding problem involves a tradeoff between two conflicting phenomena. The first involves the instrument. As the spatial resolution is increased (i. e., the angular field of view made smaller), the amount of energy received by a given telescope detector in a given spectral interval is decreased. Thus to maintain the same statistical observational error, the instrument must look at the region of interest for a longer period of time. In order to make the observational time as short as possible, the instrument should have as low a spatial resolution as possible.

The second conflicting phenomenon involves the clouds themselves. As shown earlier, decreasing the spatial resolution (i. e., increasing the IFOV) decreases the percent of area which is seen as cloudfree. In many cases the cloudfree area goes to zero because there are no holes in the clouds larger than the field of view. Thus the spatial resolution should be as high as possible. Since this latter phenomenon is more significant the greater the cloud cover, it must first be decided for what cloud cover a sounding instrument

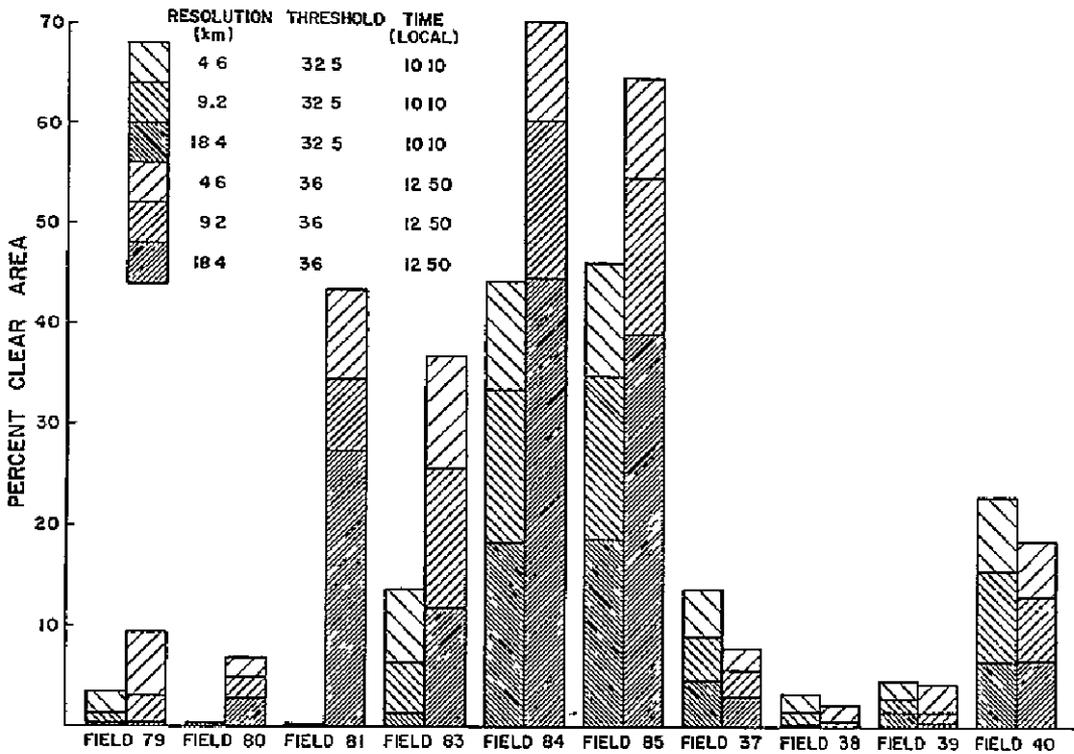


Fig. 13. Change in percent clear area in a few hours for a subtropical high pressure zone cumulus and stratus (Fields 79-85) and the intertropical convergence zone (Fields 37-40). January 20, 1968 (162940Z) and (191131Z).

should be designed. Arking *et al.** are working with averages over cloudy and clear areas to determine the probability of finding clear areas in specified geographic regions as a function of the size of the clear area. The first areas studied have a rather large percentage clear area (69% on the average). The Wisconsin study deals only with cloudy areas since the soundings of greatest interest will be near meteorologically active situations (i. e., clouds), and if one can sound near clouds, certainly one can sound in clear areas. In other words, a sounder should be designed for cloudy regions.

A sounding can be made by averaging the signals of all the cloudfree instantaneous fields of view within a given area. (The GARP report recommends

*Personal communication based on unpublished report: "Cloud Statistics Affecting the Sounding of the Lower Atmosphere from a Satellite. Preliminary Report: July 25, 1969" by Albert Arking, Frank Weinstein and Jay Fleischman. Institute for Space Studies, GSFC, NASA, New York, New York.

an area 400 km on a side.) Numerical models require the data to represent the large scale circulation; thus the input must be characteristic of the total area employed. The technique discussed in earlier sections fulfills this demand by considering many measurements over a relatively large area. Since the time base of the large scale variations is several days, it seems reasonable, if necessary, to spend several hours making a sounding. Thus the earth can be scanned several times to gather the required number of radiance measurements. Temperature soundings inferred from such measurements will be asynoptic. However, they will also provide time derivatives of the temperature profile, data not obtained from a standard radiosonde network. The inherent time and space averaging must be done in conjunction with data interpolation and smoothing required by numerical models. (GARP Report, 1969.) Another channel with higher resolution having point by point correlation with the sounding channels (perhaps an infrared mapping channel) would be required to determine whether a given instantaneous field of view is cloudfree for the reasons stated earlier (i. e., the contrast in results between Methods A and B).

TABLE 2

Number of grids which can be sounded in a given time at a given resolution. The first 18 grids of Table 1 are used.

Resolution (km)	< 1 Hour	< 2 Hours	< 4 Hours	< 6 Hours
9.2	2	5	9	11
13.8	9	10	13	14
18.4	10	11	11	11
23.0	9	9	9	9
27.6	7	7	7	7

Table 2 which gives the number of "soundable" grids in a given time at a given resolution assumes a camera in geosynchronous orbit with 16 inch optics, spinning at 100 rpm and stepping 0.2 mrad each revolution. It is looking at a spectral band 20 cm^{-1} wide centered at $14 \mu\text{m}$ with a detector having $D^* \approx 2 \times 10^{10} \text{ cm Hz}^{1/2} \text{ watt}^{-1}$. The camera scans the whole earth enough times to give an error of $0.25 \text{ erg per cm}^2 \text{ sec ster cm}^{-1}$ over an area 400 km on a side at nadir (Krauss, 1969) in the areas of interest, namely those presented in Table 1. This type of instrument was studied because it can be retrofitted to the camera system proposed for the Synchronous Meteorological Satellite (SMS). Table 2 shows that a temperature sounding radiometer is detector noise limited at smaller IFOV and "cloud noise" limited at larger IFOV. The times in this table can be decreased if only that part of the earth which contains dense clouds is scanned repeatedly. The conclusion to be drawn from this sparse data is that the ideal resolution for sounding in cloudy regions from geosynchronous orbit is somewhere between 13.8 km and 18.4 km.

7. Summary

This study has examined the percentage of cloudfree area within defined regions. Furthermore, this parameter (percent clear area) was studied in relation to a varying IFOV of a possible satellite sensor. In addition, regions including various kinds of meteorological (i. e., cloud) conditions were studied at two times $2 \frac{3}{4}$ hours apart.

It is difficult to generalize the results from this relatively small sample; more work of this kind is needed. However, one result seems evident: the percent clear area in a region decreases as the spatial resolution of the sensor is decreased. This result is expected. In addition, the rate of decrease of the measured parameter is increased as cloud "contaminants" smaller than the IFOV of the instrument are considered.

These preliminary results, based on actual measurements, from one of the highest resolution satellite sensors available to date (the ATS Multicolor Spin-Scan Cloud Camera), may be interpreted and used for various applications. In the present study, meteorologically active regions have purposely been chosen as areas of interest. This choice was made since we are interested in assessing the ability of an infrared sounder on a geosynchronous satellite to obtain vertical temperature profiles in and near these areas of prime meteorological interest. These sparse data, when combined with other studies of instrument design and accuracy to optimize an infrared sounding system, indicate that a good spatial resolution at nadir for a sounder is about 15 km. This is equivalent to about 0.4 mrad angular resolution from geosynchronous orbit.

Of general interest is the indication that clouds smaller than the instrument IFOV are very important. This indicates that determining a radiance level threshold in a visible channel is not sufficient by itself to decide whether small clouds are within the instrument's IFOV. Studies similar to the present paper should use simultaneous near infrared and thermal infrared data to see if some combination of threshold criteria in several spectral channels can detect cloud contamination with better precision. Furthermore, the sensor used to detect cloud effects should have higher spatial resolution than any sounding instrument.

Acknowledgments

We thank Mr. Eric Smith for developing the computer programs; Mr. Robert Krauss, Dr. Kirby Hanson, and Professor Verner Suomi for their helpful discussions; and Miss Betty Erwin for typing the manuscript. The study was supported by NASA, through its Laboratory for Atmospheric and Biological Sciences.

References

- COSPAR Working Group VI Report to IOC, 1969: Systems Possibilities for an Early GARP Experiment, p. 50.
- GARP Report, 1969: Plan for U. S. Participation in the Global Atmospheric Research Program. National Academy of Sciences, p. 16-17.
- Hanson, Kirby J., 1969: Applications for Bispectral Radiance Measurements from a Satellite. Ph.D. Thesis, Department of Meteorology, University of Wisconsin.
- Krauss, Robert, 1969: Performance Analysis of a Proposed High Accuracy Geostationary Sounding and Imaging System. Annual Report on Contract NAS 5 11542, in Press. Available from Space Science and Engineering Center, 1225 West Dayton Street, Madison, Wisconsin 53706.
- Salomonson, Vincent V., 1969: Cloud Statistics in Earth Resources Technology Satellite (ERTS) Mission Planning. GSFC preprint X-622-69-386. Available from Goddard Space Flight Center, Greenbelt, Maryland.
- Smith, William L., 1968: An Improved Method for Calculating Tropospheric Temperature and Moisture from Satellite Radiometer Measurements, Mon. Wea. Rev., 96, 387-396.
- Thomsen, R., R. Parent, and V. Suomi, 1969: ATS-I Spin Scan Camera. Weather Motions from Space, ed. by V. Suomi and K. Hanson, Madison, University of Wisconsin Press, in press.
- Vonder Haar, T. H., 1969: Meteorological Applications of Reflected Radiance Measurements from ATS-I and ATS-III. J. Geophys. Res., 74, 5404-5412.

N71-11621

SATELLITE MEASUREMENTS OF THE EARTH'S RADIATION BUDGET DURING A FIVE-YEAR PERIOD¹

Thomas H. Vonder Haar

ABSTRACT

This paper summarizes an extended time series of measurements of the earth's radiation budget from the first and second generation United States meteorological satellites. Values of planetary albedo, infrared radiant emittance and the resulting net radiation budget are now available for more than 35 months during the period 1962 to 1966. These measurements show a mean global albedo of 30%, and net radiation balance within measurement accuracy. The discussion treats global and zonally averaged values for the "mean annual" case, for "mean seasons" and includes a comparison of measurements during the same seasons in different years. The role of these radiation budget measurements in the total global energy balance is noted.

1. Introduction

Earth orbiting satellites provide an ideal platform for measuring the energy exchange between earth and space. Until this platform became available about ten years ago, our knowledge of this energy exchange was based exclusively on theoretical and empirical calculations. The need for the information is fundamental, since the equator-to-pole gradient of energy transfer across the "top of the atmosphere" is the prime forcing function that drives our atmospheric and oceanic circulations. Measurements from satellites now allow us to study and understand our radiation budget. Regarding the future of our atmosphere, such measurements will allow us to consider the effects of natural or inadvertent changes within our atmosphere on the global radiation budget, and thus on global weather and climate.

This paper presents a synopsis of measurements from the 1st generation (TIROS-type) satellites, together with the most recent radiation budget measurements from our 2nd generation (NIMBUS and ESSA) spacecraft. Only experiments that have measured both components (the infrared and reflected solar) of the

¹Paper presented at the Symposium on the Future of the Atmosphere, Madison, Wisconsin, October, 1969. Also presented (with V. Suomi) at the AMS/RMS Conference on the Global Circulation of the Atmosphere, London, August, 1969.

radiation budget are included (see Table 1). The first of these were obtained from TIROS-IV in 1962 (House (1965)). The most recent data available are preliminary results from ESSA-III in late 1966, early 1967 (MacDonald, 1969). Within this period, Vonder Haar (1968) presented measurements obtained during 33 months during 1962-1965. (Part of his data were collected by Bandeen et al. (1965).) Raschke and Bandeen (1970) provide a detailed discussion of two and one-half months of measurements from NIMBUS-II during 1966.

All of these data are considered in the present study. However, nearly eighty percent of all observations thus far have been acquired by the lower resolution, Wisconsin-type sensors (Suomi et al., 1967), the remainder from medium resolution scanning radiometers. A summary of results obtained through 1965 is discussed by Vonder Haar and Suomi (1969).

TABLE 1
Available Radiation Budget Measurements

Year							
Season	1962	1963	1964	1965	1966		
MAM	x		x	x		→	<u>MAM</u>
JJA		x	x	x	x	→	<u>JJA</u>
SON		x	x	x		→	<u>SON</u>
DJF		x	x		x	→	<u>DJF</u>
							↓
							<u>ANNUAL</u>

NOTE: The DJF (December, January and February) measurements are plotted under the year for the respective December. Arrows indicate the method of averaging to form mean seasonal and annual values.

2. Radiation Budget Measurements; The Mean Annual Time Period

2.1 Global and Hemispheric Results

Figure 1 shows the "mean annual" average of the radiation budget parameters integrated over the entire earth and separately for the Northern and Southern hemispheres. The parameters are related by:

"MEAN ANNUAL" RADIATION BUDGET

	PLANETARY ALBEDO			INFRARED LOSS			NET RADIATION		
	N H	S H	GLOBAL	N H	S H	GLOBAL	N H	S H	GLOBAL
1st GENERATION SATELLITES [MAM, 1962; JUNE 1963 - NOV 1965] 33 MONTHS, 1962-1965]	29	29	29	33	33	33	+ 01	+ 01	+ 01
1st GENERATION PLUS NIMBUS-II AND ESSA-III [39 MONTHS, 1962-1966]	30	30	30	.33	34	34	00	.00	00
	PERCENT			CAL CM ⁻² MIN ⁻¹			CAL CM ⁻² MIN ⁻¹		

SOLAR CONSTANT = 195 CAL CM⁻² MIN⁻¹

Fig. 1. Mean annual radiation budgets of the entire earth-atmosphere system, the northern hemisphere and the southern hemisphere as measured from the 1st generation meteorological satellites together with first results from the 2nd generation spacecraft.

$$RN = \left(\frac{\pi R^2}{4\pi R^2}\right) I_0(1.0 - A) - W_L \quad (1)$$

with RN, the net Radiation; R, the earth's radius; I_0 , the solar constant at mean earth-sun distance ($1.95 \text{ cal} \cdot \text{cm}^{-2} \cdot \text{min}^{-1}$); A, the planetary albedo; and W_L , the infrared radiant emittance from the earth-atmosphere system.

The figure shows that both the 1st and 2nd generation satellite measurements present a consistent picture of the mean annual global and hemispheric radiation budgets. Absolute measurement accuracy is estimated to be one unit of the least significant digits shown. Using mean values of the parameters in equation (1) yields relative accuracy for this time and space scale of 2 - 3%.

A summary of both the individual and combined data sets in Fig. 1 indicates:

- a) Over a 4 to 5 year period the net radiation budget of the entire earth, and of each hemisphere separately, is in radiative balance as well as we can measure such a balance. With both hemispheres in balance, there is no requirement for a net energy exchange across the equator.
- b) The entire earth and both hemispheres are darker ($A = 30\%$ vs. 35%) and warmer ($W_L = 0.34$ vs. 0.325) than earlier estimates had shown (London, 1957). The equivalent blackbody temperature difference is 3°K (254° vs. 251°). Being both warmer and darker indicates that the earth-atmosphere system must accommodate (and most probably transport) more energy than previously believed ($\sim 15\%$ more in each hemisphere).

- c) Each hemisphere has nearly the same planetary albedo and infrared loss to space on the "mean annual" time scale. This points out the dominant influence of clouds on the energy exchange with space, since the surface features of the two hemispheres are quite different.

2.2 Zonal Averages and Geographical Variations

Figure 2 continues our view of the "mean annual" case. It shows the measured meridional profiles of planetary albedo, absorbed solar energy and infrared loss. The classic picture of excess energy absorbed at low latitudes and net energy loss from midlatitudes and polar areas is apparent. Note, however the lack of equatorial symmetry, with more energy retained in the southern subtropics than at the same latitudes in the north. This occurrence is offset by a difference in the polar radiation budgets. The Arctic is observed to be warmer and darker than the Antarctic. Latitudinal averages of planetary albedo are as low as 23% in the subtropics, ranging to 50% near the north pole and 70% at the south pole. These polar values are heavily weighted by the summer-time conditions.

With the aid of Fig. 3 we take a closer look at the radiation budgets in two latitude zones of approximately equal area. For the regions 0 - 10°N and 60 - 90°N, the "mean seasonal" energy gain or loss measured by satellites is shown by the bar graph, the "mean annual" average by the solid line. London's annual value for each zone is noted by a dashed line. On this space scale, smaller than that of the globe or a hemisphere, we find a wide range of departure between measurements and earlier computations. For the region 0 - 10°N we find that the net annual gain of energy calculated by London is 35% less than that measured from satellites. However, for 60 - 90°N, London's value falls much closer to the observed mean. Recall that the measurements have shown (and earlier calculations require) near hemispheric balance of gains and losses on this time scale. This occurs because zones not shown in Fig. 3, especially at mid-latitudes, have a measured net loss exceeding that computed (primarily to the higher values of infrared loss already mentioned).

The departure of measurements from early estimates in the deep tropics is of special significance. Astronomical factors insure large amounts of available solar energy during the entire year. The large energy gain shown in Fig. 3 results primarily from a lower planetary albedo (Fig. 2). Thus, the tropics gain more energy because they are "darker." More than any other single factor, this result of satellite measurements contributes especially to the total global albedo being lower than previously believed. In addition, the greater energy input at low latitudes implies either greater poleward energy transport by ocean currents or greater air-sea exchange of energy and increased tropical convection. Vonder Haar and Hanson (1969) used the satellite measurements together with a summary of the few available observations of solar radiation received at the

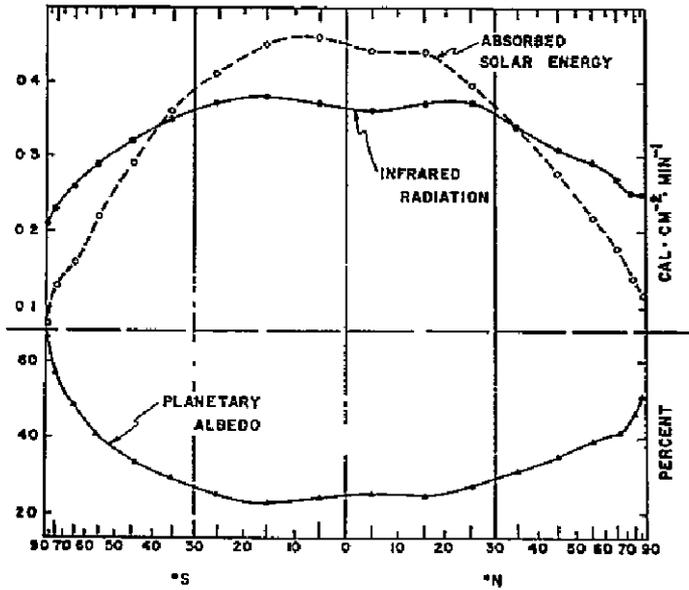


Fig. 2. Mean meridional profiles (averages within latitude zones) of components of the earth's radiation budget measured during the period 1962-1966. The abscissa is scaled by the sine of latitude.

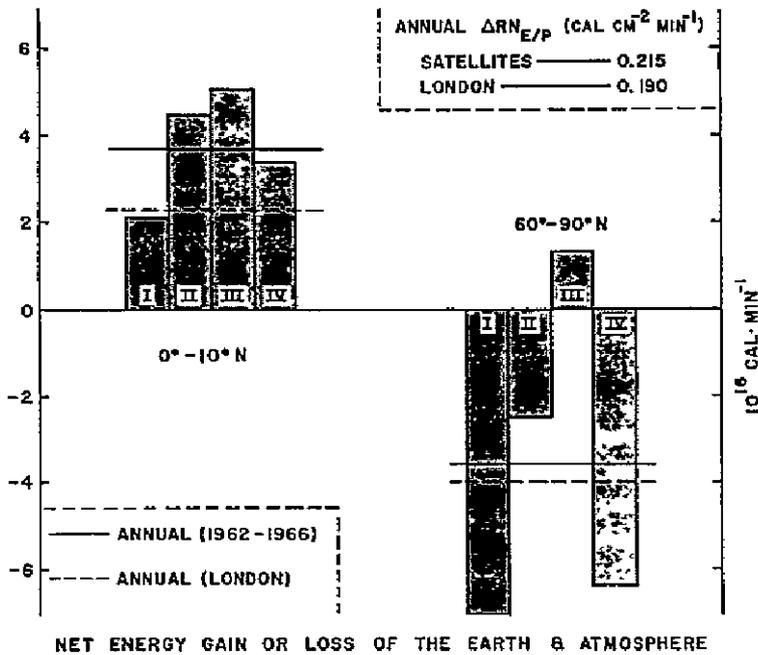


Fig. 3. Mean annual and seasonal energy exchange with space, measured from satellites during 1962-1966, for two latitude zones. Bar graph represents seasonal values (I = Dec., Jan., Feb., etc). $\Delta RN_{E/P}$ is the net radiation gradient between equator and pole.

surface in the tropics (Quinn and Burt (1967)). The surface data show that 20% more energy arrives annually (in the region 0 - 20°N) than was earlier estimated from climatic data and empirical relations (i. e., Budyko (1963)). Thus, the combination of recent measurements are in agreement and show that additional energy entering the earth-atmosphere system is primarily added to the oceans.

Figure 3 also shows the wide seasonal range of net energy exchange with space over the Arctic cap. From a large net loss in the winter, to a small but significant net gain during the summer season, the range is about two and one-half times greater than that for 0 - 10°N. Figure 7 will show that the southern polar regions do not have this wide a seasonal range, pointing out again the lack of equatorial symmetry in the energy exchange with space.

An insert in Fig. 3 notes that the "mean annual" pole-to-equator gradient of net radiation measured from satellites is 10 - 15% greater than London's mean annual value for the Northern Hemisphere. This simple index is a measure of the planetary forcing function, the required poleward energy transport by the atmosphere and oceans. It indicates that our circulations are more "vigorous" than previously estimated.

A closer look at time variations measured over the north polar cap is presented in Fig. 4. It shows a 17 month time series of the incident, absorbed and reflected solar energy, the infrared loss, albedo and net radiation. The earliest satellites (TIROS) had orbital inclinations such that they did not view this polar region. Thus, Fig. 4 shows our first long period of radiation measurements over the Arctic. Second generation NIMBUS and ESSA satellites, in near polar orbits, also returned measurements over the poles, but during shorter time intervals of our five-year period (Raschke, Bandeen and Möller (1967)). An interesting feature in Fig. 4 shows that the absorbed solar energy maximum lags a month behind peak insolation. Several earlier studies had anticipated this observation based on information about the break-up and puddling on the Arctic ice pack during the summer. Even though the infrared loss reaches a flat maxima in July, the total radiation budget of the earth-atmosphere system is slightly positive. The time period of net gain began in May and ends abruptly in August; it occurs primarily because of the features of the solar energy budget.

Thus far we have only discussed global and hemispheric averages and the mean radiation budget within latitudinal zones. This can be misleading, because an important result of the satellite measurements shows that significant changes in the earth's radiation budget occur within latitude zones, especially in the tropics. Figure 5 presents mean annual maps of the radiation data. These geographical variations of planetary albedo infrared radiation and net radiation will be discussed elsewhere in more detail. Note, however, the zonal pattern of isolines at high latitudes in both hemispheres. This occurs in the region of migratory storms, where alternating clear and cloudy conditions

occur over a given area. In the northern hemisphere a continental influence can be seen, especially in the map of planetary albedo. In the tropics semi-permanent features of the atmospheric circulation, terrain features, or the influence of special conditions in a specific season (i. e., the monsoon) cause significant departures within the latitude zones. This is true even in the map of net radiation. Recall the maps are for the mean annual case, and thus all daily and even seasonal anomalies are smoothed a great deal. Even so, we see a definite distribution of relative energy gain and loss areas with a zone. In higher latitudes (more zonal patterns) the range of variation is much smaller. These results point out that a numerical model to simulate the circulation of the atmosphere or ocean cannot add a forcing function only as a function of latitude over longer time periods (i. e., as a function of solar declination). Such an input may be valid for the higher latitudes but certainly not for the tropics. Variation in net radiation is shown in Fig. 6 for three tropical zones. Some of the greatest minima are found over the oceanic deserts west of South America and Africa. Here the low, bright, warm clouds reflect the solar energy well and also radiate strongly in the infrared. Maxima of net radiation occur over clear oceanic regions.

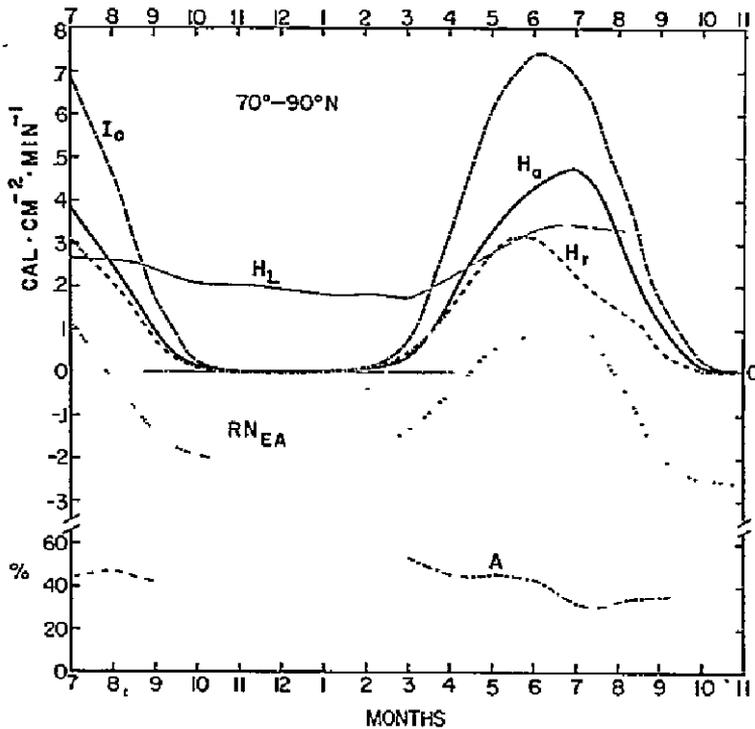
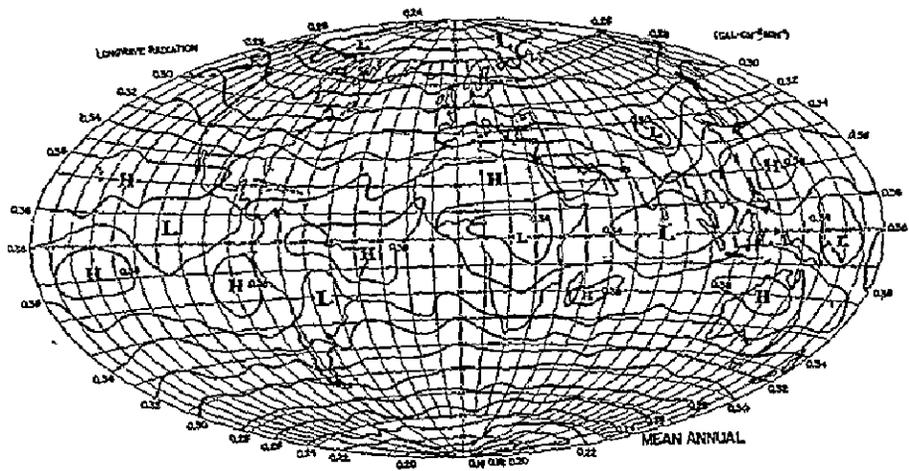


Fig. 4. Continuous time series of radiation budget parameters over the north polar region during 1964 and 1965. H_a and H_r are the portions of solar radiation absorbed and reflected by the earth-atmosphere system. I_0 is the mean monthly insolation computed with a solar constant of $2.00 \text{ cal} \cdot \text{cm}^{-2} \cdot \text{min}^{-1}$.



(a)

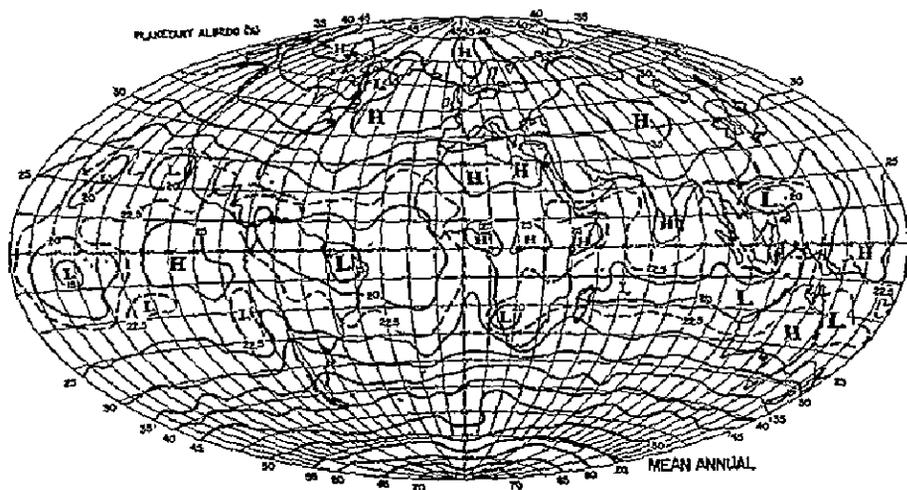


Fig. 5. Example of the geographical distribution of radiation budget parameters. Planetary albedo and longwave radiation measured from satellites during 1962 - 1965,

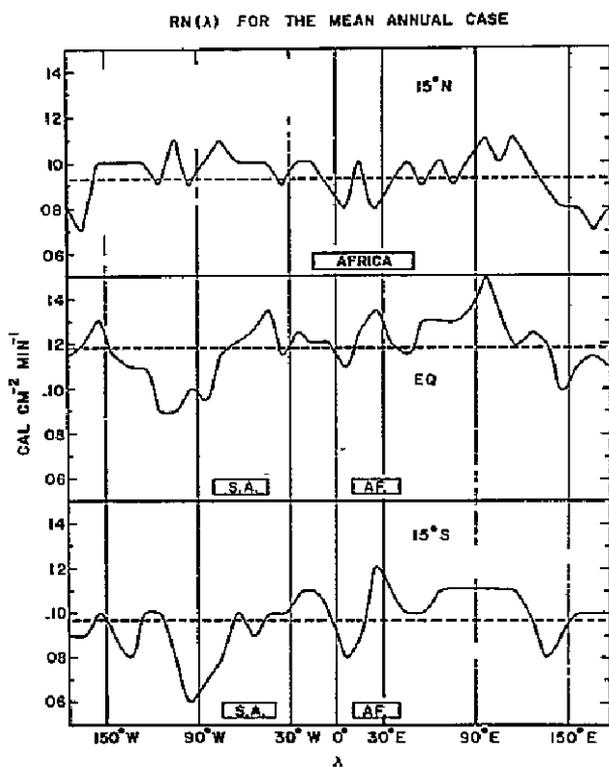
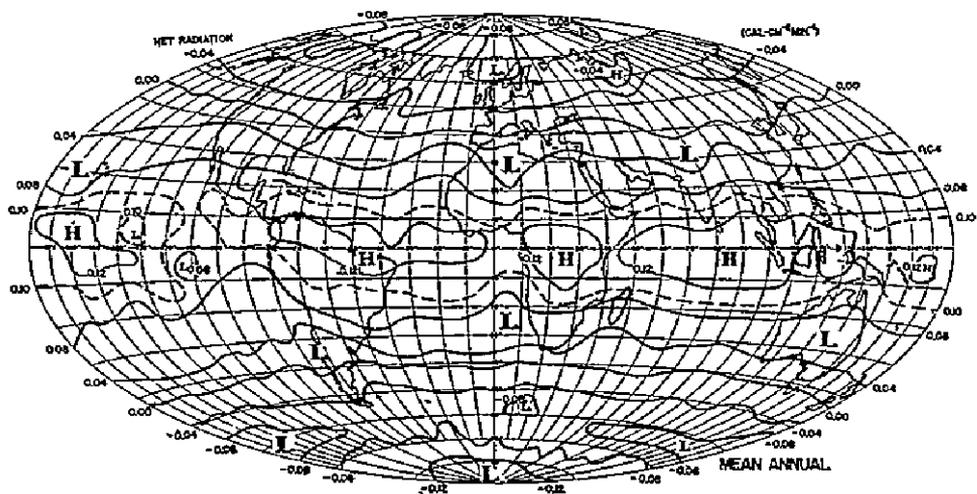


Fig. 6. The variation of net radiation with longitude (λ) extracted from Fig. 5 for three latitude circles. Dashed line is the zonal mean value; major land masses are shown schematically.

3. Mean Seasonal Budgets and Interannual Variations

Reference to mean and specific seasonal measurements has been included above for comparison with the mean annual case. Figure 7 gives a more detailed view of the average zonal net radiation as it varies with latitude during each season of the year. As in Fig. 2, we note the departure from exact symmetry about the equator. For example, in the north polar region has the wider range of net radiation with season. The net radiation reversal near the winter pole results from the absence of solar radiation at the highest latitudes combined with increasing infrared loss to space from the relatively warmer regions away from the poles.

For a basic understanding of the energy exchange between earth and space, mean annual and seasonal averages provide the first step. Now that a relatively large sample of measurements is available we can proceed further, and examine observed interannual variations of the radiation budget. We use a simple index ($\Delta R_{NE/p}$) the difference in net radiation between the equator and $90^\circ N$ or S. In Fig. 8 we see the change in the measured index during the same season of different years. The horizontal line shows the mean annual gradient, slightly

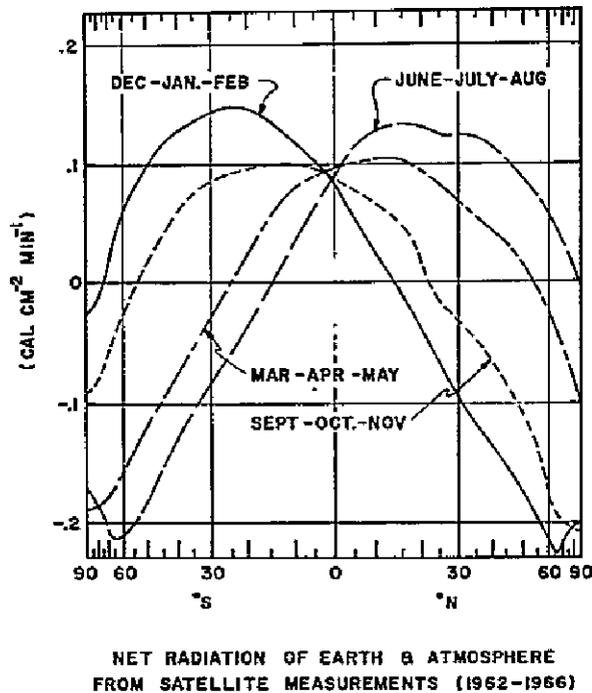


Fig. 7. Mean seasonal values of the meridional profiles of net radiation, obtained from satellite measurements and using a solar constant of $1.95 \text{ ly} \cdot \text{min}^{-1}$.

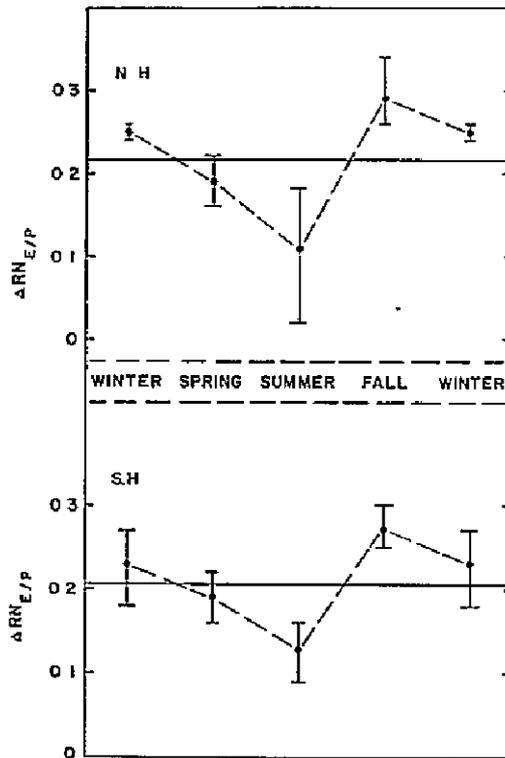


Fig. 8. Mean annual and seasonal values of the equator-to-pole gradient of net radiation and the range of interannual variation measured from satellites. Table 1 shows the periods of measurement; units are $\text{cal} \cdot \text{cm}^{-2} \cdot \text{min}^{-1}$. Both northern and southern hemisphere results are shown.

larger in the northern hemisphere. Dots indicate mean seasonal radiation gradients and the bars note the range of values during the three or four years of data for each season.

We observed the largest gradient change between summer and fall in both hemispheres; a change toward a stronger gradient, a stronger forcing function. The greatest range within one season was observed during the northern hemisphere summer, the smallest range during winter of the same hemisphere. Absolute magnitudes of the gradient are largest during fall. However, the reversal of net radiation seen in Fig. 7 is not considered by our simple index. An upward adjustment for winter would make that gradient nearly the same as for fall. As expected, summer gradients are small in both hemispheres.

A further step will be to study the response of atmosphere to each specific season's measured gradient of net radiation. The measure of the forcing function used in this study (or another similar index) can now be obtained routinely by our meteorological satellites. Of course, we must consider the large thermal inertia of our oceans and thus expect the response and feedback to vary with both time and location. Some initial work in this area has been reported by Winston (1967).

4. The Role of Radiation Budget in the Global Energy Balance

The energy exchange between the earth and space by radiative transfer is only one component of the total energy budget of the earth-atmosphere system. True, on a global scale and over long time periods it is the overwhelmingly significant component, but we are very often concerned with higher frequency variations in space and time. In this section we shall use measurements of the earth's radiation budget together with estimates of the other major terms in order to present a simple depiction of the earth's total energy budget. We do this to illustrate the role of the radiation terms in the energy budget. Since much weight must be placed on climatological estimates of certain parameters the absolute accuracy of results in this section is open to question. In addition, the illustrative exercise falls short of a thorough discussion of atmospheric energetics, the study of generation, conversion and dissipation of various "forms" of energy in the atmosphere.

Figure 9 shows, for the mean annual case, the total poleward energy transport (RT) required by the measured radiation budget. Vonder Haar (1968) and others have shown that for no net energy storage in the atmosphere, earth or ocean the net required energy transport across a parallel of latitude must equal the sum of several horizontal energy transport terms. These are: ΔF , energy transport by ocean currents; LAC_y , the transport of energy in latent form as water vapor; and ΔC , the sensible heat plus potential energy transport by the atmosphere. To derive Fig. 9 we used estimates of ΔF and LAC_y from a recent compilation by Sellers (1965). Thus, using satellite measurements for RT we obtain ΔC as a residual. Note that values plotted above the zero line in Fig. 9 indicate northward energy transport, below the line—southward. We note a double maxima of ΔC in both hemispheres, very similar in shape to one derived by Holopainen (1965) from northern hemisphere radiosonde data. Note also the requirement for southward transport of sensible heat and potential energy across the equator, but northward movement of water vapor by the atmosphere.

It is possible to derive similar curves for the mean seasons (Rasool and Prabbakara (1966)). However, a major uncertainty in the transport terms arises from lack of knowledge about seasonal variation of energy storage in the oceans. It may be expected that within some latitude zones, during some seasons, that the radiative, latent heat (i. e., evaporation minus precipitation), and

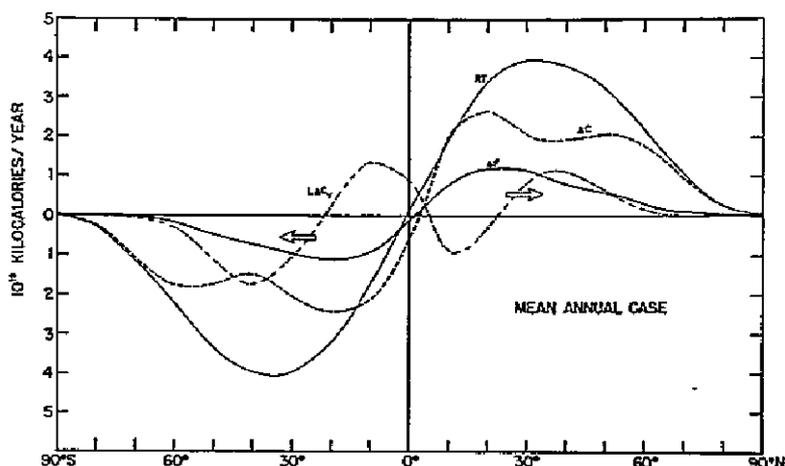


Fig. 9. Total poleward energy transport required by the radiation budget (RT) and its major components.

ocean storage terms may all be nearly the same magnitude. Newell *et al.* (1970) discuss new measurements of the atmospheric transport terms. Of course, satellite measurements will be available on a continuing basis. We may eventually be able to study oceanic transport and storage as a residual.

5. Conclusions

A summary of this first comprehensive study of measurements of the earth's radiation budget shows that:

- a) The 1st generation meteorological satellite observations as well as the early results from our second generation spacecraft have independently shown a consistent picture of the earth's radiation budget. It is a warmer and darker planet than we previously believed. More solar energy is being absorbed, primarily in the tropics.
- b) Although the tropics ($30^{\circ}\text{N} - 30^{\circ}\text{S}$) as a whole gain energy from space during all seasons, significant longitudinal variations in the solar input are noted.
- c) Although the annual net radiation budgets of the northern and southern hemispheres are both in balance (despite the difference in topography) we do not find exact equatorial symmetry in radiative exchange with space. This is especially true in subtropical and polar regions.

- d) The prime forcing function, the equator-to-pole gradient of net radiation, has its greatest relative change between summer and fall in both hemispheres. In addition, there have been significant gradient changes measured between the same seasons in different years. Further study of atmospheric response is in order.
- e) In like manner, even though the satellite measurements show no requirement for net energy exchange across the equator for the "mean annual" case, the measurements can be combined with independent observations to gain a better understanding of the global energy balance and the transport requirements of the atmospheres and oceans.

Further Comments

The results of satellite measurements shown in this paper are the first of their kind. Like all new observations, final judgment of their relevance must rest primarily on confirmation by independent means. The first such confirmation is already on record, coming from sensors flown on the first of our second generation satellites. Of course, the absolute and relative accuracies of the measurements can be improved in future sensing systems. See, for example, error analyses included in the references dealing with the Wisconsin-type sensors (since these provide 80% of the measurements reported here).

The trend toward departure from the earlier computation studies of the radiation budget seems irreversible. The reason for departure from previous estimates is still not completely resolved, although we strongly suspect that an overestimate of tropical cloudiness, especially as it affected solar energy transfer, must have been a major failing of the early studies. Indeed, London¹ has re-evaluated his 1957 study with new cloud information and computes a global planetary albedo of 31%. Thus, the computations are beginning to converge on the observations. The end result of computations which use satellite measurements as a control at the upper boundary should be improved estimates of the vertical profiles of net radiative cooling in the atmosphere. Such profiles are very important in atmospheric energetics studies and cannot be easily measured on a global scale.

Throughout this paper we have used London's 1957 study as representative of a great deal of extensive earlier computational work. It serves this purpose well and allows us to avoid repetitive comparison of the satellite measurements with the multitude of previous numbers computed by a like number of methods. In the same vein, to maintain some fluidity in this paper we have only

¹Personal communication.

discussed the great number of satellite measurements collectively via their time and space means. A great deal of research remains when the measurements of the radiation exchange between earth and space are applied to special studies. Our initial work on interannual variations is a small start in this direction. Indeed, further work on these kinds of problems, not possible before, may be of more scientific interest than a polished description of mean conditions.

Acknowledgments

We thank our many colleagues at the University of Wisconsin who have worked on the satellite radiation budget experiments, especially Professor Robert Parent and Mr. Eric Smith. The support of NASA, through Contracts NASw-65 and NAS5-11542 from their Laboratory of Atmospheric Sciences, Goddard Space Flight Center, has been the backbone of our work during a ten-year period. In recent years, scientists at the National Environmental Satellite Center of ESSA have aided our work as well.

References

- Bandeem, W. R., M. Halev, and I. Strange, 1965: A radiation climatology in the visible and infrared from the TIROS meteorological satellites, NASA TN D-2534, 30 pp.
- Budyko, M. I. (Ed.), 1963: Atlas of the heat balance of the globe (in Russian), Moscow, 69 pp.
- Holopainen, E. O., 1965: On the role of mean meridional circulations in the energy balance of the atmosphere, TELLUS, XVII, 3, 225-294.
- House, F. B., 1965: The radiation balance of the earth from a satellite, Ph. D. Thesis, Department of Meteorology, University of Wisconsin, 69 pp.
- London, J., 1957: A study of the atmospheric heat balance. Final Report on Contract AF 19(122)-165, Department of Meteorology and Oceanography, New York University, 99 pp.
- MacDonald, T. H., 1969: National Environmental Satellite Center, ESSA (personal communication).
- Newell, R. E., et al., 1970: The energy balance of the global atmosphere (to be published in Quarterly Journal of the Royal Meteorological Society).

- Quinn, W. H., and W. V. Burt, 1967: Weather and solar radiation reception in the equatorial trough, J. Appl. Meteor., 6, 988-993.
- Raschke, E. and W. R. Bandeen, 1970: The radiation balance of the planet earth from radiation measurements of the satellite NIMBUS II, J. Appl. Meteor., 8, 2.
- Raschke, E., F. Möller and W. R. Bandeen, 1968: The radiation balance of the earth-atmosphere system over both polar regions obtained from radiation measurements of the NIMBUS II meteorological satellite. Scientific Paper Dedicated to Dr. Anders Angstrom, Medellanden, Serie B, Nr. 28, Severiges Meteorologiska och Hydrologiska Institute, Stockholm, 104 pp.
- Rasool, S. I. and C. Prabhakara, 1966: Heat budget of the southern hemisphere pp. 76-92 in Problems of Atmospheric Circulation, ed. by Garcia and Malone, Spartan Books, Washington, D. C., 186 pp.
- Sellers, W. D., 1965: Physical Climatology. The University of Chicago Press, Chicago and London, 272 pp.
- Suomi, V. E., K. J. Hanson and T. H. Vonder Haar, 1967: The theoretical basis for low-resolution radiometer measurements from a satellite, Annual Report, Grant WBG-27, Department of Meteorology, University of Wisconsin, 79-100.
- Vonder Haar, T. H., 1968: Variations of the earth's radiation budget, Ph. D. Thesis, Department of Meteorology, University of Wisconsin, 118 pp.
- Vonder Haar, T. H., and K. J. Hanson, 1969: Absorption of solar radiation in tropical regions, J. Atmos. Sci. 26, 4, 652-655.
- Vonder Haar, T. H. and V. E. Suomi, 1969: Satellite observations of the earth's radiation budget, Science, 1963, No. 3868 (14 Feb.), 667-669.
- Winston, J. S., 1967: Zonal and meridional analysis of 5-day averaged long-wave radiation data from TIROS-IV over the Pacific sector in relation to the northern hemisphere circulation, J. Appl. Met., 6, 3, 453-463.

N71-11622

A STUDY OF THE INDIAN MONSOON USING SATELLITE
MEASURED ALBEDO AND LONG WAVE RADIATION

Gerald J. Dittberner

ABSTRACT

Thirty months of continuous low resolution satellite data, emitted long wave radiation and albedo are used to study the Indian monsoon. These data, combined with rainfall data for the Indian subcontinent, are employed to develop an intensity index which provides the first quantitative measure of monsoon intensity. Three monsoons are studied: the 1963 monsoon, the strong 1964 monsoon and the very weak 1965 monsoon. Intensity differences show up most dramatically on the intensity index maps. The time history of the 1964 and 1965 monsoons are examined showing how the intense cloudiness of the 1964 monsoon lingered on for a number of months apparently reducing surface heating significantly preceding the weak 1965 monsoon. Typical values of the intensity index, long wave radiation and albedo are compiled for each season of the year preceding the 1964 and 1965 monsoons. Finally, the monsoon is described in terms of radiation depicting certain recurring, quasi-steady state, seasonal features.

CONTENTS

1. Introduction	256
2. Historical Background	256
3. Data Availability	258
4. Procedure of the Investigation	259
4.1 Long Wave Radiation, Albedo and Rainfall Relationships	261
4.2 Intensity Index Nomogram	263
4.3 Test of the Intensity Index	264
5. The Indian Monsoon Using the Intensity Index	265
5.1 The History of Two Monsoons	270
5.2 Description of the Monsoon in terms of Radiation	276

6. Summary and Suggestions for Further Research	278
7. References	280

1. Introduction

The term "monsoon," according to Webster, is "a periodic wind, especially in the Indian Ocean and southern Asia; also, the rainy season of the southwest monsoon in India." The word is derived from the obsolete Arabic "Mausim," the Dutch "Monsson," and the Portuguese "Mancao" meaning "a time, a season."

Predicting the monsoon onset is still one of the most challenging problems in meteorology. Rainfall during the southwest monsoon season (June to September) is extremely important to the agriculture of India and the economy of the country (Pant (1964) and Shamshad (1966)). In fact, 90% of the population of India is agricultural (Raman (1963)).

To date there have been many studies of the monsoon, some of which are described in section 2. Most of these investigations were restricted to the use of data collected from land stations, with very little being available from the vast ocean areas. New data has recently become available. Vonder Haar (1968) in his study of the earth's radiation budget, accumulated more than 30 months of continuous, satellite measured albedo and radiated long wave data over almost the entire globe.

These data encompass the period from the summer of 1963 to the autumn of 1965. Thus, for the first time, continuous radiation data over land and sea are available which cover the growth and development of two monsoons. Ironically, the monsoon of 1964 was very strong and the monsoon of 1965 was very weak and practically nonexistent.

The main task, then, is to examine these two monsoons over a longer span of time and space than has ever been attempted before. Further, characteristic radiation patterns of the monsoon will be suggested. In the process, there appear indications that the long wave radiation data can be combined with the albedo yielding a quantitative measure of monsoon intensity.

2. Historical Background

Differential heating as the cause of the monsoon was set forth as early as 1686 by Halley in the Memoirs of the Philosophical Society of Great Britain. This theory is still unchallenged.

Simpson (1921) states that the primary cause of the monsoon is the high

temperature and corresponding low pressure over the land. Air is then transported under the influence of this pressure system and the rotation of the earth from the southern hemisphere northwest across the equator and then northeast in the northern hemisphere. The moisture accumulated in the long journey over the oceans is then distributed over India due to orographic effects in and around India. Thus, he says, the rainfall associated with the monsoon is determined by topography, and says nothing about other causes.

Wagner (1931) held that the monsoon represents a stationary system of cyclonic disturbances reaching up to the average height of the Himalayas between two air masses, one continental belonging to the west winds of the middle latitudes and the other maritime. He considers that the continental air is warmer than the monsoon air at the surface and colder above and that a considerable part of the monsoon precipitation in and south of the Gangetic plain is due to cyclonic convergence.

The Army Air Force in their study of India, Burma and Southern China (Climate and Weather of Southeastern Asia, 1942) carries this effect further observing that these Eastern cyclones move toward the low over northern India during the monsoon season and then move southward in late September. They conclude that the pulse and extent of the monsoon are related to these depressions.

Yin (1949), in his renowned paper on the 1946 monsoon, discovered that the burst of the monsoon occurs as a mean low latitude trough that is displaced rapidly from one semipermanent position near 90 E to another relatively semi-permanent position near 80 E. He finds that the motion of this trough is caused mainly by the displacement of the low level westerly jet from south of the Himalayas to their northern boundary. This displacement correlates in time with a general rearrangement of the northern hemispheric long wave pattern that results in a replacement of a mean ridge by a mean trough over central Siberia. A polar trough then extends all the way from Siberia to the tropics.

Chakravorty and Basu (1956) investigated the influences of disturbances from the west on the weather over northeastern India during the monsoon season. They state that the contribution of such disturbances when they pass through the eastern Himalayas is to increase rainfall over the submountain regions of northeast India to a significant and often important degree. They maintain that the passage of such disturbances accentuates the monsoon trough near the foot of the Himalayas. They also state that the monsoon current becomes stronger due to the convergence of the Arabian Sea and Bay of Bengal branches of the monsoon.

Pant (1964) extended the work of Yin (1949) by incorporating the changes in large scale circulation as related to the gradual onset of monsoon rain over different parts of India. He maintains that the monsoon trough forms near 90 E

at 700 mb and then shifts westward to 80 - 85 E and becomes more intense. The orientation of the primary axis of the monsoon trough, which was north-south at the beginning of June, becomes east-west by the end of June, the monsoon low having formed with its axis along about 22 N. This is the time, he says, when the monsoon establishes itself over the whole country. At the same time, the Pacific High shifts northward toward Japan. He concludes that the onset of the monsoon thus occurs gradually and calls for new methods to forecast the formation and movement of the monsoon trough. It is worth noting here that the availability of satellite photographs and radiation data now provide a substantial data base for developing such forecast methods.

Colon (1964) studied the Arabian Sea area and interactions between the southwest monsoon current and the sea surface. He states that over large portions of the Arabian Sea there is a rapid warming of the surface water during late winter and early spring. The maximum temperatures are observed around May, at the time of the establishment of the southwest monsoon circulation. There is then a cooling trend to a minimum in August and September. The water cooling appears to be a direct result of the establishment of the southwest monsoon regime. Large rates of heat flux by evaporation (670 langleys per day or 1.2 cm per day) were observed in the west central portions of the Arabian Sea. He shows that the evaporation makes a major contribution to the water cooling in that area.

Rao (1966) used long wave radiation (8 - 12 microns) from TIROS IV to study the 1962 monsoon. He observes that the centers of low out-going long wave radiation associated with the cloudiness of the intertropical convergence zone (ITC) moved northward from April to June, the most abrupt shift occurring in the middle of May. With the advance of cloudiness, the monsoon sets in over the Indian subcontinent. Changes in the radiation values showed changes in the cloud fields. He also noted that when more satellite data become available it will be possible to study the variations in the monsoon circulation, the retreat and the development of the monsoon by observing the changes in the radiation intensities.

The large number of investigations point out that the monsoon is still not completely understood. First generation satellites (the TIROS series) have provided the first good indication of what is happening over the vast ocean areas surrounding the Indian subcontinent. As satellite data increases in quality and quantity our understanding of the monsoon will likewise increase, thus yielding more effective methods of forecasting the onset, intensity and duration.

3. Data Availability

Radiation data used here are essentially those of Vonder Haar (1968). Most of the data were collected by low resolution sensors whose design was based on principles first described by Suomi (1958). Data for June, 1963 to

September, 1963 were collected by hemispheric sensors mounted on TIROS IV. These are described in detail by Sparkman (1964) and House (1965). The data were transformed to radiation budget parameters (long wave radiation, albedo, and net radiation) using the techniques developed by House (1965) and Suomi et al., (1967).

Data for June, 1963 to May, 1964 were collected from medium resolution sensors mounted on TIROS VII. A complete description of this system is given by NASA Staff Members (1964). Seasonal averages of albedo and long wave radiation were computed in a manner similar to that of Bandeen, et al. (1965).

Data for June, 1964 to November, 1965 were collected by disc sensors mounted on experimental satellites. The description and data reduction techniques are similar to those for the hemispheric sensors, above.

The data is given in terms of an average value for each 10×10 degree latitude-longitude block since the low resolution sensors require this limit on spatial resolution. Long wave radiation is given in langleys per minute and albedo is given in percent (representing 0.3 - 3.0 microns, see Vonder Haar (1968)).

The change of sensors in June, 1964 would lead one to suspect comparisons in any time series. However, there were about three months of overlapping data. These were compared by Vonder Haar and found to correlate almost exactly.

The rainfall data were compiled from the Indian Journal of Meteorology and Geophysics (for example: Staff Members (1964)). The data are given in terms of percent of mean seasonal rainfall for the 30 meteorological divisions of India.

It should be noted that because of the agricultural aspect of the monsoon, rainfall data in percent are used rather than actual millimeters of rainfall. That is, the most important consideration to the economy of India is whether the monsoon rains were above normal or below normal.

4. Procedure of the Investigation

Long wave radiation and albedo at the top of the atmosphere are generally indicative of cloud top heights and cloud amounts, respectively. Fritz and Winston (1962), Wark, Yamamoto, and Lienesch (1962), and Rao and Winston (1963) have shown methods of converting long wave radiation values to cloud top heights. Although there are difficulties when dealing with high thin cirrus or scattered clouds, it is assumed that values of 0.34 langleys per minute or below correspond, generally, to broken or overcast clouds and that regions of

lowest outgoing radiation correspond to the thickest clouds or clouds of large vertical depth (Rao (1966)).

High values of albedo are indicative of large cloud amounts. The average albedo of the earth is about 30%. Vonder Haar (1968) calculated a value of 29% (± 1). The average cloud cover of the earth is 0.4 to 0.5. Thus about half the earth is clear to scattered and half is broken to overcast. The oceans, which cover most of the earth, under clear conditions have an albedo of about 10% (Conover 1965) and taking into account some scattered clouds and land, the average albedo for the "clear" half of the earth will be about 15%. This means that the average albedo for the "cloudy" half is about 45%. To include broken situations, a value of 28% or above is chosen to represent regions with broken to overcast conditions.

The study of Rao (1966) shows how long wave radiation can be used to study the monsoon. Similar studies could be conducted using albedo. Each type of data is adequate for examining the southwest monsoon (and other meteorological phenomena). But using both types in conjunction with one another should provide much more insight.

In this light, an attempt is made to combine three parameters (long wave radiation, albedo, and rainfall) into an index representing the intensity of the monsoon.

It is generally accepted that more precipitation is expected from clouds with large vertical extents than from low clouds and that more precipitation is expected from areas with high cloud amounts than from areas with low cloud amounts. By identifying regions with low long wave radiation (high cloud tops) and high albedo (large cloud amounts), areas of maximum cloud activity can be located.

Figure 1 shows this region qualitatively. That portion with low long wave radiation and high albedo is labeled high, thick, broken or overcast. The region with high long wave and high albedo represent bright, warm targets such as fog or low cloud decks. The region with high long wave radiation and low albedo represent warm, generally nonreflecting targets such as cloudless oceans. The region with low long wave radiation and low albedo represent cold, generally nonreflecting targets or a combination of high, thin cirrus with some radiation emanating from lower layers or the earth's surface coming through.

A point located region I is expected to represent greater activity than a point in any of the other regions.

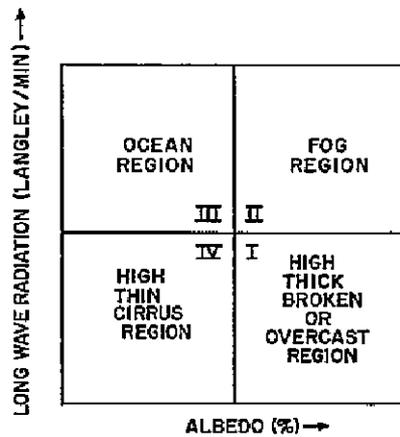


Fig. 1. Qualitative diagram of long wave radiation vs. albedo showing regions explained in the text.

4.1 Long Wave, Albedo and Rainfall Relationships

The Indian southwest monsoons for 1964 and 1965 provide an excellent opportunity to investigate the relationships between radiation data and rainfall because of the striking differences in intensity. In fact, there was enough difference in the two to cause the annual total agricultural production level to suffer a sharp drop in 1965 as shown in Fig. 2 (based on data from Food and Agriculture Organization of the United Nations, 1968).

An average precipitation percents for the three 10 X 10 degree blocks, which approximate the Indian subcontinent, were computed taking into account the different areal coverage of each meteorological division. Points corresponding to region I of Fig. 1 appear to show the best correlation. A diagram of these vs. the appropriate long wave radiation value (in time and space) is shown in Fig. 3. A similar diagram with albedo is shown in Fig. 4. These points are quite scattered, as expected, since actual rainfall is not a function of long wave radiation and/or albedo along but is affected by topography, local convection, etc. There does, however, seem to be some relationship.

The distribution of points in Figs. 3 and 4 is different. The distribution for albedo vs. rainfall is wider than that for long wave radiation vs. rainfall; this implies that actual rainfall is more dependent on cloud top height and thickness than on cloud amounts, as expected.

For the purpose of using these results as the first approximation to an intensity index, the distribution is considered to be the same on the two diagrams, and linear. The results of this investigation will show that this

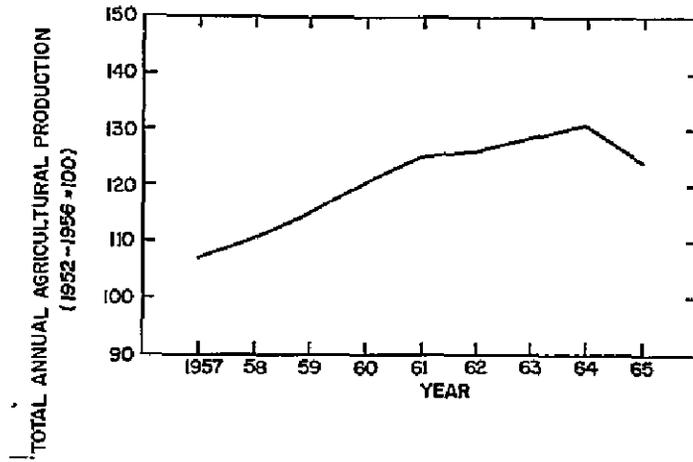


Fig. 2. Total annual agricultural production of India for 1957 through 1965.

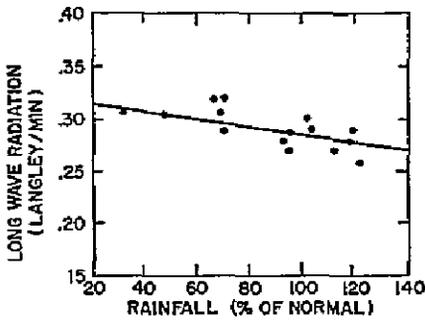


Fig. 3. Long wave radiation vs. rainfall.

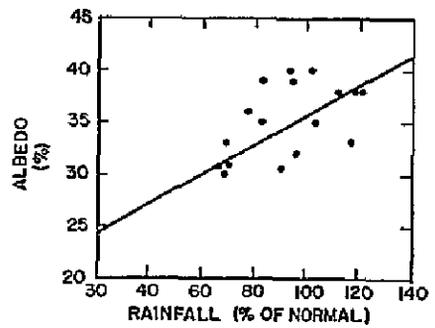


Fig. 4. Albedo vs. rainfall.

first approximation does, in fact, produce meaningful results, thereby calling for further investigations to determine an optimum index.

4.2 Intensity Index Nomogram

With the results of the previous section, a nomogram is constructed in the fashion of Fig. 1. The nomogram is constructed by taking points at unit distances for each value of rainfall from the relationships in Figs. 3 and 4 and plotting these on a diagram of long wave radiation vs. albedo, as shown in Fig. 5. The rainfall value of each point is taken as the index number that is shown on Fig. 5.

The nomogram can be interpreted in several ways. For example: given values for the long wave radiation and albedo for a certain region, the nomogram can be entered and the intensity index read off directly. The index represents a measure of meteorological activity with respect to clouds and an indication of intensity. Higher index values mean higher activity or higher intensity.

Lines of constant intensity index can be interpreted as follows: If the two radiation values are given, and the corresponding index determined, a certain cloud intensity is defined. Now, if there is another value of long wave radiation which is slightly lower, with the same albedo, more cloud activity is expected since this lower value represents higher cloud tops. In the same way, given an albedo which is slightly larger, with the same long wave radiation, more cloud activity is expected since this larger value represents a larger cloud amount.

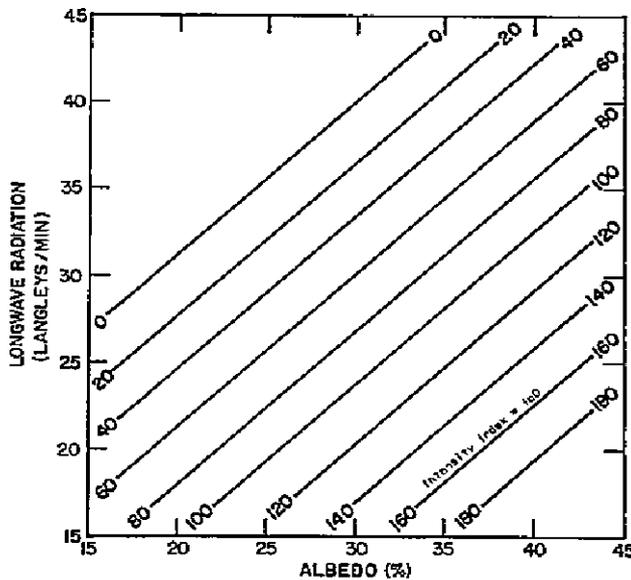


Fig. 5. Intensity Index Nomogram

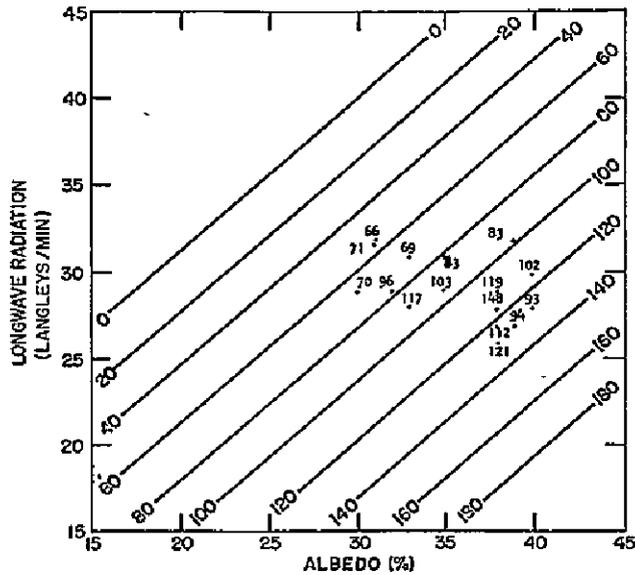


Fig. 6. Intensity index nomogram with actual precipitation (% of normal).

It must be noted that great care is to be used when trying to apply this index over land surfaces. Misleading index values can occur, for example, because of high albedos from clear deserts (such as the Sahara), or low long wave radiation values over snow or ice, especially in high and middle latitudes in winter.

It is also expected that this nomogram (which is still a first approximation) will work best in the region of the graph representing low long wave radiation and high albedo.

4.3 Test of the Intensity Index

On Fig. 6 the nomogram is reproduced and has, additionally, the actual precipitation values for the points falling in region I of Fig. 1. Note that most of the actual rainfall data fall reasonably close to the corresponding intensity index. There also appears to be some indication that for large actual precipitation values the intensity index is too low. Perhaps one reason for this is that actual precipitation data is being compared to the empirical intensity index. In addition, the derived index is only a first approximation. Even so, the data fits well enough to suggest that refinement of the index would be a worthwhile effort.

Now that a certain degree of confidence in the index has been reached, its use shall be extended to study the Indian southwest monsoon not only over

land, but out over the vast data sparse region of the oceans surrounding the Indian subcontinent.

5. The Indian Monsoon using the Intensity Index

In this chapter, the Indian monsoon is studied using long wave radiation, albedo and the intensity index. It is shown that either long wave radiation or albedo can be used to study the monsoon but can also produce misleading results. It is also shown that the intensity index combines the best features of both types of data producing meaningful, quantitative results.

Seasonal maps of long wave radiation, albedo and the intensity index are presented in Figs. 7, 8 and 9. The season depicted encompasses June, July and August (JJA) and defines the monsoon season as used in this paper.

Figure 7 depicts the monsoon seasons of 1963, 1964 and 1965 in terms of long wave radiation. Shaded areas denote regions with values of 0.34 langley's per minute or less, and generally represent broken to overcast conditions as described previously. In the monsoon season of 1963 (Fig. 7), two areas of high clouds dominate. The main area is over the southeast Asian peninsula and a smaller area covers the ocean southwest of Kerala (on the southern tip of India). Clear conditions exist over the Arabian peninsula and eastward to about 70 E. Another clear area is oriented along 17 S. Some cloudiness also is in evidence over Africa at about 10 N.

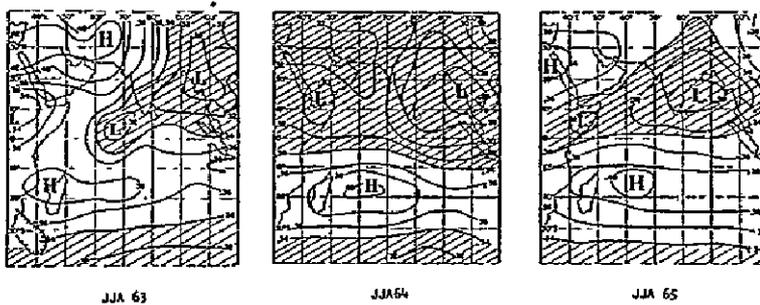


Fig. 7. Mean seasonal values of outgoing long wave radiation during the monsoon season (June, July and August) for 1963, 1964 and 1965. Units are in langley's/min. Shaded regions represent values less than 0.34 langley's/min.

The intertropical convergence zone (ITC) can be seen situated along 10 to 15 N, as evidenced by the regions of lowest radiation in the monsoon pattern and the low values over Africa.

The monsoon cloudiness has a minimum value of 0.30 langley's per minute northeast of the Bay of Bengal and extends from the equator to at least 40 N to the east of India. The western boundary is approximately along a line from central Tibet to the equatorial Arabian Sea. The seasonal cloudiness covers all of southern India and most of the northeast portion. Note that while the seasonal mean shows little activity in the northwest, an analysis of the monthly data indicates that the monsoon was most active in the northwest states in August and September. Such activity is not expected to show up on the JJA averages.

The cloudy area southwest of the peninsula is interesting since it approximates the location of high net sea-air heat flux mentioned by Colon (1964).

The intense monsoon season of 1964, as shown in Fig. 7, depicts extensive cloudiness covering the entire portion of the map north of the equator. The African radiation minimum is farther north than in 1963 and is centered over Aden. Cloudiness covers the Arabian peninsula which was clear the previous year. South of the equator, the clear area is more evident but is still in the same location.

In addition to covering a larger area, the monsoon cloudiness is more intense. In fact, the minimum radiation is only 0.25 langley's per minute compared to 0.30 in 1963. The radiation values over India are lower by about 0.07 langley's per minute.

The ITC is slightly north of its 1963 position over Africa and slightly south in the Bay of Bengal.

There is very little evidence of a cloudy region southwest of the peninsula on this map.

In 1965 the monsoon was very weak, especially in the northern states. But Fig. 7 shows the region of cloudiness to be larger than in 1963. This apparent paradox will be considered later when the albedo is taken into account in determining intensity indices.

In 1965, then, the cloudiness appears to cover almost all of India, southeast Asia, the Arabian Sea and the Bay of Bengal. The double minimum is still present although the main minimum is shifted farther west by about 10 degrees. The southern hemisphere clear area is in the same position but is extremely zonal compared to the previous two years.

The two radiation minimums have the same values as in 1963 although the pattern appears to be more widely separated. Note that the ITC is located south of its 1963 position over Africa by about 10 degrees.

Again, there is little evidence of a cloudy region southwest of the peninsula.

Considering the three seasons depicted in Fig. 7, the general features of the monsoon (such as the center over the Bay of Bengal) appear each year. Using these three maps alone, one can observe that the 1964 monsoon was indeed the strongest. However, it is not clearly evident that the 1965 monsoon was the weakest! Before proceeding on this point, the albedo maps merit examination.

Figure 8 portrays these three monsoon seasons in terms of albedo. Shaded areas signify regions with values of 28% or more, and generally represent areas with broken or overcast conditions as previously described.

The monsoon cloudiness of the 1963 season (Fig. 8) extends from Kerala eastward over the Bay of Bengal and southeast Asia and northward to at least 40 N. The highest albedo value of the monsoon is 37%. High albedo values also exist over the Arabian peninsula and north Africa. Note that the Arabian and Sahara deserts alone have albedo values near 30% (Conover, 1965). This fact combined with the high long wave radiation values in Fig. 7 indicate an area that is almost completely clear. The large values near 10 N over Africa, however, are due to cloudiness since this area is not a desert region.

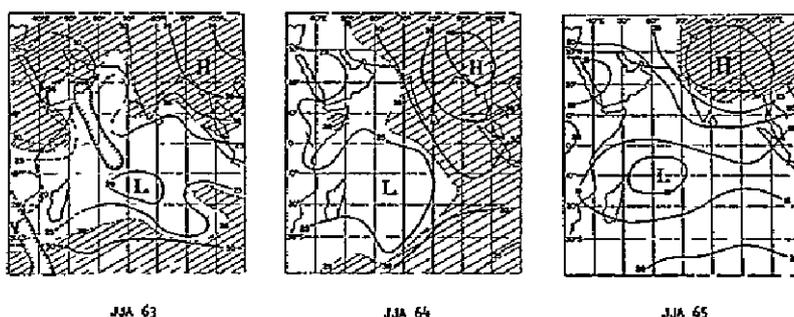


Fig. 8. Mean seasonal values of planetary albedo during the monsoon seasons of 1963, 1964 and 1965. Units are in percent. Shaded regions represent values greater than 28%.

Some evidence exists supporting the existence of the cloudy region southwest of the peninsula. The clear area south of the equator is well defined by low albedo values.

The strong monsoon of 1964, in Fig. 8, again shows cloudiness covering the entire northeast quadrant of the map, and almost all of the eastern half. The highest albedo is now 43% compared to 37% in 1963 and is in the same position. The cloudiness over the Somali Republic appears as a small region of high albedo values. Southwest of the peninsula, there does appear some evidence of a cloud region which was not apparent on the long wave radiation map (Fig. 7).

The clear area in the southern hemisphere is again broader than in 1963.

The 1965 monsoon season (Fig. 8) shows cloudiness north and east of central India. There is little evidence of broken or overcast skies on any other portion of this map. The clear area in the southern hemisphere is very pronounced, as is shown by albedo values of 10%.

Recall that the albedo of oceans with no clouds is about 10% (Conover, 1965).

The albedo maximum is shifted westward about 10 degrees, which corresponds to the westward shift on Fig. 7.

Once again, considering the three seasons on Fig. 8, the general features do appear on all three maps but not as impressively as before. Using these three maps alone, it is evident that 1964 was the strong monsoon year and 1965 was the weak year.

It is clear, from the preceding discussion that either long wave radiation or albedo alone can be used to describe, in part, the Indian monsoon. But neither tells the whole story; long wave radiation is lacking in its ability to determine cloud amounts and is perhaps misleading by itself. Albedo maps are lacking in the ability to locate the ITC, for example, and is misleading in areas with bright backgrounds.

Figure 9 presents the three monsoon seasons in terms of the intensity index derived in section 3. Shaded areas denote regions with index values of 40 or more and represent zones with the most intense activity.

In the monsoon season of 1963 (Fig. 9) areas of high activity cover southern India, the Bay of Bengal and southwest Asia including northeast India and Tibet. The zone of high activity over northern Africa shows up very well, with index values up to 47. The activity southwest of the peninsula is only moderate with an index of 28. In the strongest portion of the monsoon,

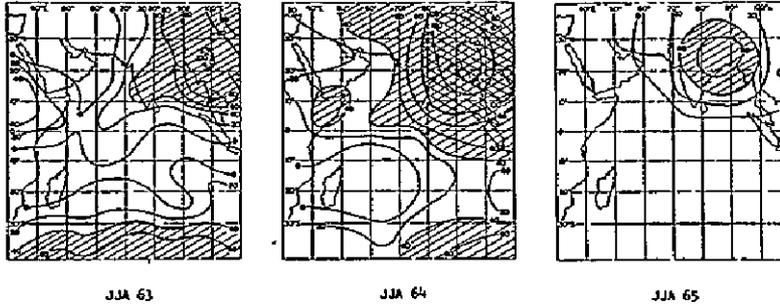


Fig. 9. Mean seasonal values of nondimensional intensity index for the monsoon seasons of 1963, 1964 and 1965. Shaded regions represent values greater than 40 and cross-hatched regions represent values greater than 100.

centered over eastern Burma and western Laos, activity is very intense with a maximum index of 100. Here one sees a true picture of monsoon activity depicted in terms of a quantitative measure of intensity.

The strong monsoon of 1964 portrays itself vividly on Fig. 9. The most intense portion of the monsoon covers more than the northeast quadrant of the map and has index values up to 155 over Burma. High values exist throughout India, the Bay of Bengal and southeast Asia. The extension southwest of the peninsula is quite pronounced and active with index values up to 50. Note that the southern hemisphere clear area has index values of zero and larger in extent than that in 1963.

In 1965 (Fig. 9) the maximum monsoon activity index is less than 70, and very restricted in size. It does not quite cover India, Tibet and Burma, and covers only a portion of the Bay of Bengal. There is no other activity on this map.

The apparent paradox, mentioned earlier, was that the long wave radiation map for 1965 (Fig. 7) indicated the same, if not more cloudiness than the map for 1963 when, in fact, the 1965 monsoon was much weaker. The reason comes to light by examining the corresponding albedo maps (Fig. 8). The albedo values over the Arabian Sea in 1963 were around 24% while those in 1965 were only around 17%. When these are combined into index values (Fig. 9), this difference is readily apparent. Thus another limitation on using long wave radiation by itself has been found. It is clearly obvious from Fig. 9 that the 1964 monsoon was the most intense and the 1965 monsoon was the

weakest. Further, by observing the maximum index values for each year, it can be determined quantitatively, how much more intense the 1964 monsoon was compared to, for example, the 1963 monsoon. The maximum index for 1963 is 100, for 1964, 155, and for 1965, 70.

Before one can say that one monsoon is, say, one and one-half times as intense as another, it is necessary to refine this index into its most optimum form. Crude as the present index is, the results outline a method by which the intensities of various meteorological phenomena can be investigated.

5.1 The Time History of Two Monsoons

Thirty months of continuous radiation data provide an excellent opportunity to examine monsoon development, progression and withdrawal. Time-latitude and time-longitude sections are chosen as the best method of investigation. The time-latitude sections depict average values for each 10 degree latitude belt from 40 N to 40 S and 30 E to 110 E. Similarly, the time-longitude sections portray average values for each 10 degree longitude belt from 30 E to 110 E and from 40 N to 20 S.

Advection of cloudiness or intensity northward or southward can be detected on the time-latitude sections, and eastward or westward motion, on the time-longitude sections.

Figures 10 and 11 show long wave radiation in terms of time-latitude and time-longitude sections, respectively. Shaded areas correspond to those used on previous diagrams.

In Fig. 10 during June, July and August of 1963 (JJA 63) the most active portion of the monsoon appears between the equator and 20 N. A nearly clear area moves southward from about 40 N over a period of three seasons and stabilizes near 15 N in winter followed by a slight northerly movement in spring (MAM 64). This is the period of maximum surface heating preceding the onset. The 1964 monsoon season starts abruptly in June as shown by the strong gradient between spring and summer. Note that the monsoon trough described by Yin (1949) and Pant (1964) does not appear on these diagrams. The motion of this trough occurs in a very short period of time (from the end of May to the first part of June), and compared to the time scale used here, is not expected to appear.

Cloudiness remains from 10 S to at least 40 N for almost three seasons, becoming concentrated on the equator in DJF 64/65. It then moves northward throughout MAM 65 to its most northerly extent (about 20 N) becoming the 1965 monsoon. In fact, northern India (north of 20 N) was extremely dry during the monsoon. Cloudiness disappears completely in SON 65.

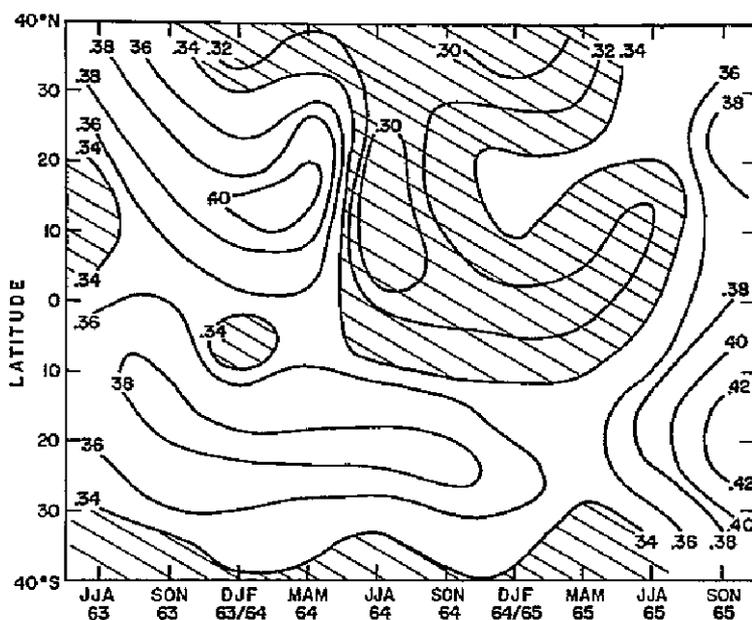


Fig. 10. Latitudinal means of long wave radiation (30 E to 110 E) in langleys/min. Values less than 0.34 langleys/min. are shaded.

Important differences in the two monsoons are now brought to light. In particular, the period of large surface heating so evident in MAM 64 was virtually nonexistent in MAM 65. The abrupt onset characterizing the 1964 monsoon occurs very gradually in 1965. Also, while the 1964 monsoon shows little influence due to cloudiness coming from the north or south, the 1965 monsoon appears to have been caused entirely by cloudiness moving northward from the equator.

The southern clear region is also interesting, moving from 10 S in the summer of 1963 to 20 S for the next three seasons and moving farther south following the 1964 monsoon. Its movement in the two years is completely different.

Figure 11 is now examined for eastward and westward movements. Very little cross-longitude motion occurs in the year preceding the 1964 monsoon. The clear zone, which moved north to south during this period, remained confined to the western region (30 E to 60 E). The abrupt monsoon onset in 1964 is also apparent on this diagram, but indicated rapid motion from the east. Cloudiness remains across almost the entire longitude range for the next three seasons. The most cloudy portion appears to move westward to about 70 E until MAM 65 when it moves eastward culminating in the 1965 monsoon.

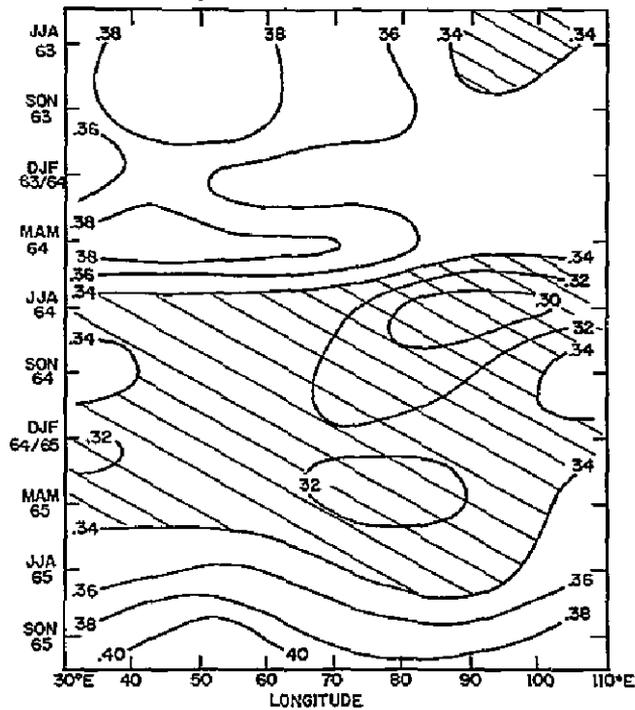


Fig. 11. Longitudinal means of long wave radiation (40 N to 20 S) in langleys/min. Values less than 0.34 langleys/min. are shaded.

Considering Figs. 10 and 11 together, it is seen that the 1963 monsoon dissipates almost completely leaving the following three seasons relatively cloud free (on a seasonal basis). During this period, surface heating takes place reaching a maximum in MAM 64, and perhaps is the main cause of the 1964 monsoon. Cloudiness from the 1964 monsoon, however, does not dissipate significantly with high clouds moving southwest to about 0°N and 75°E by the end of autumn and remaining there until MAM 65. At this point it moves northwest into the southern portion of India becoming the 1965 monsoon.

In other words, monsoon cloudiness in 1964 appears abruptly following the clear spring season. This cloudiness split into two segments, one moving northeastward, the other moving southwestward into the south Arabian Sea. More importantly, the Arabian Sea segment appears to have maintained its intensity until the 1965 monsoon season when it gradually moved into southern India. This is apparently the major source of cloudiness for the 1965 monsoon.

Figures 12 and 13 show albedo in terms of time-latitude and time-longitude sections, respectively. Shaded areas correspond to those used on previous diagrams.

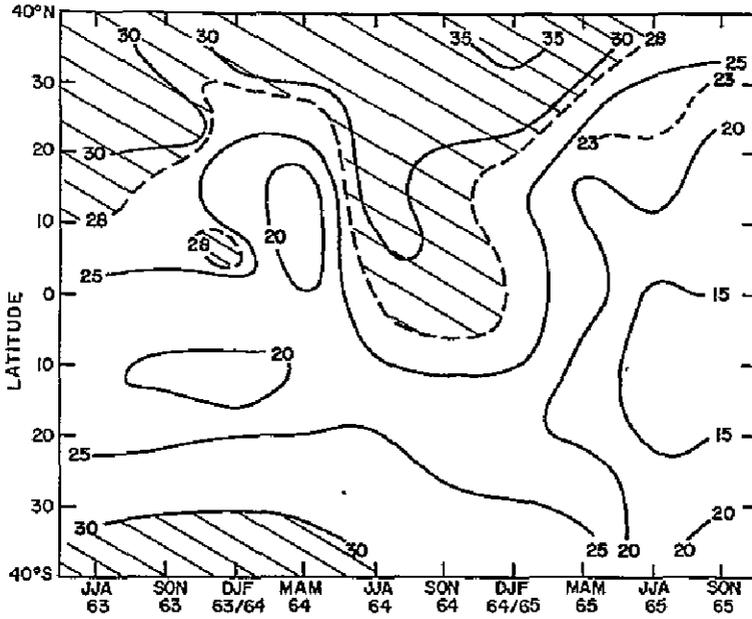


Fig. 12. Latitudinal means of albedo (30 E to 110 E) in percent. Values greater than 28% are shaded.

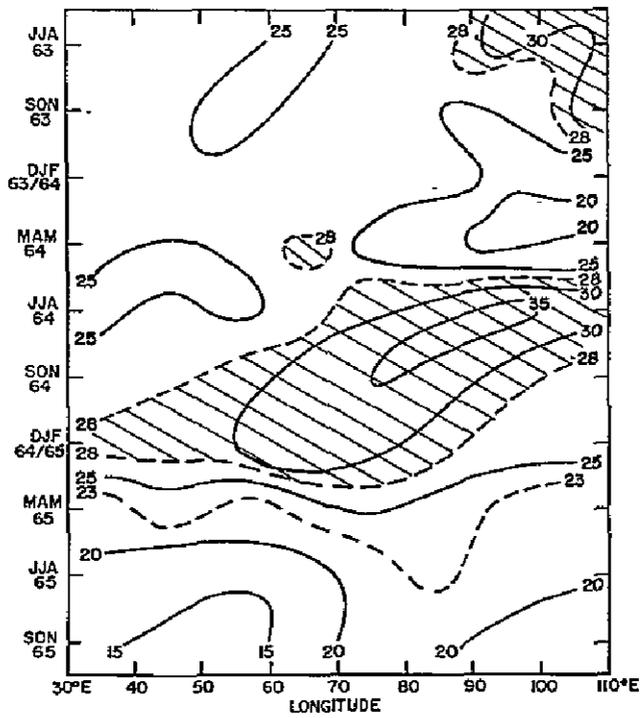


Fig. 13. Longitudinal means of albedo (40 N to 20 S) in percent. Values greater than 28% are shaded.

In Fig. 12 the apparent regions of cloud cover north of 20 N are, to a large extent, due to desert regions and must be interpreted with great care. Again, the cloudless spring of 1964 is prominent as is the abrupt increase in cloud cover associated with the onset of the 1964 monsoon. The movement of high cloud cover to the south is also apparent. A slight increase in cloud amounts appears during the 1965 monsoon.

In Fig. 13 the MAM 64 heating is apparent. The most prominent feature is associated with the post-monsoon cloud cover. It originates during the monsoon and progresses steadily westward until DJF 64/65. At this point the diagram shows a decrease in cloud cover, and that which remains moves eastward until the 1965 monsoon.

Considered together, Figs. 12 and 13 show the area of large cloud amounts dissipating after the 1963 monsoon, and the clear heating season preceding the 1964 monsoon. Again, the region of large cloud amounts following the 1964 monsoon appears. It moves southwestward until DJF 64/65. Note that cloud cover then decreases as it moves northeastward culminating in the 1965 monsoon.

Figures 14 and 15 show the intensity index on time-latitude and time-longitude sections, respectively. Shaded areas are the same as those on previous diagrams.

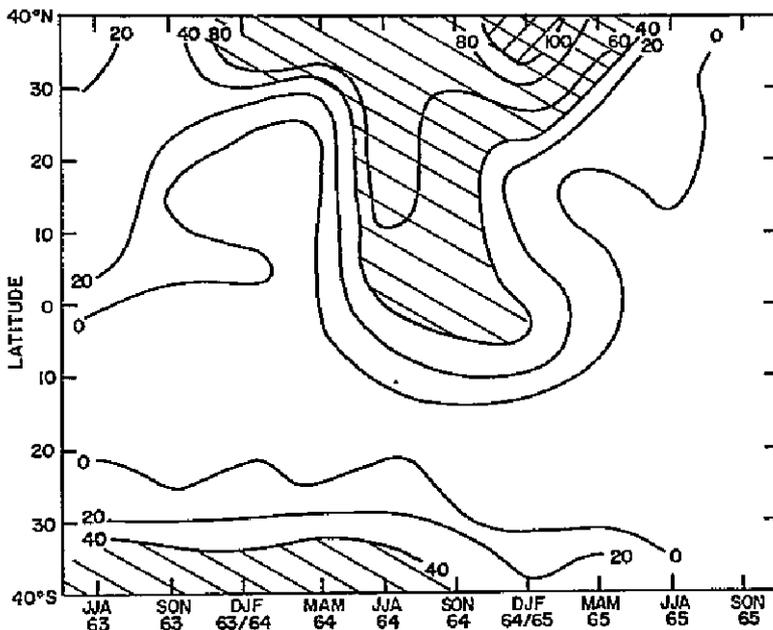


Fig. 14. Latitudinal means of nondimensional intensity index (30 E to 110 E). Values greater than 40 are shaded and values greater than 100 are cross-hatched.

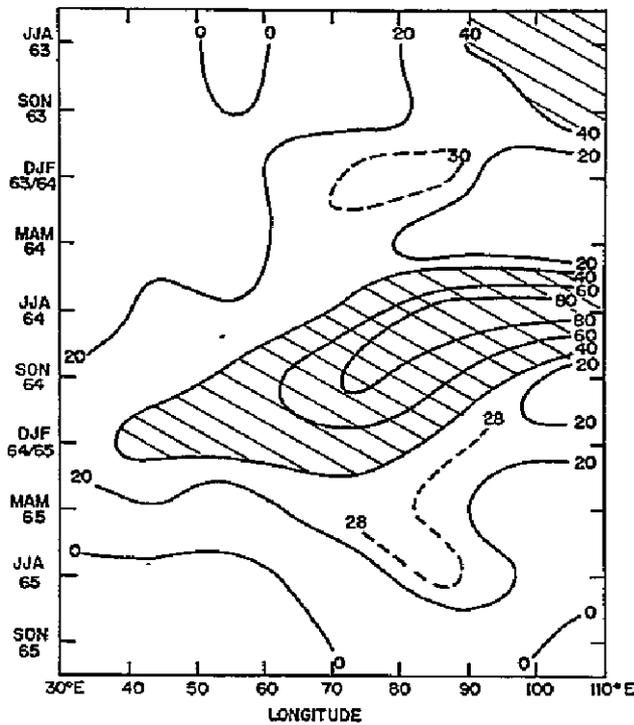


Fig. 15. Longitudinal means of nondimensional intensity index (40 N to 20 S). Values greater than 40 are shaded.

In Fig. 14 the region of greatest activity is associated with the monsoon of 1964. The southward movement throughout SON 64 into DJF 64/65 is readily apparent. Note the indication of some southwesterly motion following the 1963 monsoon. This feature is also suggested on long wave radiation and albedo sections also (Figs. 10 and 12), but its relative magnitude compared to that of the 1964 monsoon can be determined by using the index. The 1965 monsoon appears on Fig. 14 as a slight southward projection of an index isopleth.

It is noted that these values represent averages over a latitude range (30 E to 110 E) that extends well beyond the normal western extent of the monsoon. In this light, values for 60 E to 110 E were computed and are presented in Table 1. Index values from Table 1 for the Indian subcontinent (10 N to 40 N) are about 43 for the 1963 monsoon, 95 for 1964 and only 28 for 1965. These values can be interpreted as an indication of the relationship between monsoon intensity and agricultural production.

In Fig. 15 there are indications of active regions moving westward following the 1963 monsoon, but not nearly as pronounced as that following the 1964 monsoon. This intense region moves steadily westward with time and has

nearly the same latitude spread for SON 64 and DJF 64/65. Note the decrease in intensity after DJF 64/65 and the eastward motion into the 1965 monsoon.

Figures 14 and 15 together show that the active area following the 1963 monsoon moved southwestward in SON 63 decreasing in intensity, and then moved westward in DJF 63/64. But the intense region following the 1964 monsoon moved southwestward throughout SON 64 and DJF 64/65 and maintained its intensity. That is, the active area of 1964 was more intense, moved farther southwest and preserved its intensity longer than in the previous year.

There is a priori verification of extensive cloudiness throughout the tropic regions in 1965.¹

It is also instructive to consider representative values of long wave radiation, albedo and the intensity index for the seasons preceding the two monsoons. Figure 16 portrays these seasons pictorially. The thickness and width of clouds on Fig. 16 symbolically correspond to the typical values of long wave radiation and albedo shown. Values are representative of the area of the Indian subcontinent from the Himalayas to the equator. It is assumed that land areas have albedo values of about 20% and long wave radiation values of about 0.40 langley's per minute.

The intensity indices for 1964 follow the expected pattern. During the autumn (SON 63) it is quite low representing monsoon dissipation, in winter it increases slightly with the winter monsoon, in spring it is very low allowing maximum surface heating preceding the monsoon, and in the monsoon season the index is very high.

Index values do not follow the typical pattern in 1965. Monsoon activity lingers on throughout the autumn with a very high intensity index. The spring is by no means clear, as pointed out earlier, reducing surface heating significantly as shown by the low monsoon season index.

5.2 Description of the Monsoon in Terms of Radiation

Inspection of all available seasonal maps and time sections reveals certain recurring, quasi-steady state, seasonal features. Such features suggest characteristic radiation patterns for the monsoon, in terms of seasons throughout the year. These features are listed below. Discussion is limited to the area from 40 N to 40 S and from 30 E to 110 E (see Fig. 7).

The spring season (MAM): Nearly clear skies prevail from the Himalayas to 30 S all across the map. Some weak activity exists around the equator at 70 E.

¹Personal communication from Mr. Lewis Allison, NASA, based on Environmental Technical Applications Center (ETAC) cloudiness summaries.

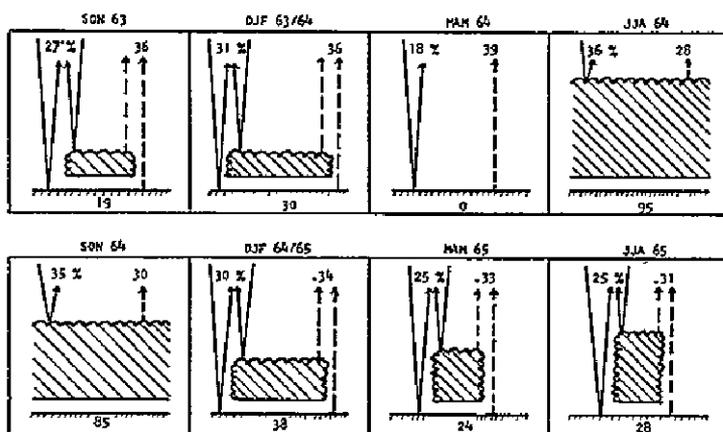


Fig. 16. Typical values of long wave radiation (dashed lines, langleys/min), albedo (solid lines, percent) and nondimensional intensity index (bottom of each season) for India and its immediate vicinity.

TABLE 1

Longitudinal Averages (60°E to 110°E)

LONG WAVE RADIATION (langleys/minute) $\times 10^{-2}$			
	JJA 63	JJA 64	JJA 65
40°N			
30°N	35.6	30.4	35.2
20°N	34.8	28.6	33.0
10°N	32.4	27.6	30.6
ALBEDO (Percent)			
	JJA 63	JJA 64	JJA 65
40°N			
30°N	32.6	36.2	28.6
20°N	33.0	36.4	29.4
10°N	29.0	34.2	23.8
INTENSITY INDEX			
	JJA 63	JJA 64	JJA 65
40°N			
30°N	40	92	24
20°N	46	99	28
10°N	42	94	22

The summer (monsoon) season (JJA): Intense activity covers India, the Bay of Bengal and southeast Asia from the equator to at least 40 N. Two centers of maximum activity make up the monsoon. The strongest is located in the region of Burma and the Bay of Bengal and the other is located in the southeast Arabian Sea near 10 N, 70 E just off the southern tip of the Indian peninsula. A large clear area exists in the southern Indian Ocean east of Madagascar. Some activity also exists over northeast Africa at 10 N.

The autumn season (SON): The two intense centers associated with the monsoon split apart. The strongest center, which was in the Bay of Bengal region, moves to the northwest into eastern Tibet and southern China. The weaker center, which was in the south Arabian Sea, moves to the southwest to a relatively stable position near the equator and 70 E. The Arabian Sea, Iran, Saudi Arabia and northeast Africa are nearly cloud free. A clear area also exists south of the equator to 30 S.

The winter season (DJF): The active region near the equator at 70 E remains in approximately the same position but becomes slightly more intense. North of 25 N cloud activity covers the entire map. These are the features usually associated with the winter monsoon.

The preliminary model discussed here is based on seasonal maps of satellite sensed radiation. It is but a suggestion of what could be done with monthly, weekly, or even daily data.

6. Summary and Suggestions for Further Research

Thirty months of continuous low resolution satellite data, emitted long wave radiation and albedo, were used to study the Indian monsoon. These data, combined with rainfall data for the Indian subcontinent were employed to develop an intensity index which provides the first quantitative measure of monsoon intensity. It was assumed that the index is also representative of precipitation over the oceans where rainfall data is very sparse. Thus a useful tool was developed to study the monsoon over land and the vast ocean areas surrounding the Indian subcontinent.

Three monsoons were studied using the intensity indices and radiation data: the 1963 monsoon, the strong monsoon of 1964 and the very weak monsoon of 1965. Intensity differences in the 1963 and 1965 monsoons were not readily apparent on the long wave radiation diagram. Some difference is observed on the albedo maps, but only on the intensity index diagrams does this difference appear so dramatically. More importantly, the past history of the 1964 and 1965 monsoons were examined. The pre-monsoon seasons proved to be the most interesting. The primary cause influencing the high intensity of the 1964 monsoon appears to be the intense surface heating during the preceding cloudless spring. It was noted that activity associated with the 1964

monsoon lingered on for several months in the south Arabian Sea contributing to cloudy skies in the spring of 1965. Because of the marked cloudiness in this latter season, surface heating was reduced significantly. Thus the cloudiness gradually moving in from the south Arabian Sea became only slightly more intense upon reaching the peninsula. This factor, combined with the apparent evidence that no other cloudy regions moved into India yielded a weak, almost nonexistent 1965 monsoon causing a significant drop in the annual agricultural production.

Typical values of the intensity index, long wave radiation and albedo were compiled for each season of the year preceding each monsoon. From this data, one finds that the intensity index for the autumn of 1964 was only slightly lower than that in the monsoon season.

A model of the monsoon in terms of radiation data was suggested. The most important observation was that the monsoon is composed of two main activity centers: one over Burma and the Bay of Bengal and one in the southern Arabian Sea just off the southern tip of the Indian peninsula. Following the monsoon season, the Bay of Bengal center moves north and east into eastern Tibet and southern China while the south Arabian Sea center moves southwest to a relatively permanent position near the equator and 70 E.

As the result of this study, new and expanded research areas are suggested. In particular, the intensity index stands only as a first approximation and calls for studies leading to a refined, optimum form. Such an index could be used to study phenomena other than the monsoon. A study using actual precipitation rates rather than percent of normal would be particularly valuable over the oceans. The precipitation term in energy budget studies is difficult to determine over the seas of the world and a refined index could provide the first quantitative values for this term.

Finally, the monsoon itself is still not completely understood. Further studies should be undertaken using higher resolution (spatial and spectral) radiation data and monthly, weekly or even daily radiation values. Such investigations would lend themselves to refining the model of the monsoon, and provide further insight into more effective methods of predicting the onset, intensity, duration and withdrawal of the Indian monsoon.

ACKNOWLEDGMENTS

I wish to express my gratitude to Professor F. Sechrist and Dr. T. Vonder Haar for their advice and guidance during the preparation of this thesis. I also thank Professor E. Wahl for his helpful comments and suggestions.

I also wish to thank my wife, Mary Kay, whose patience and understanding contributed immensely to this work.

The research in this thesis was supported, in part, by the National Aeronautics and Space Administration (NASA) contract NAS5-11542.

REFERENCES

- Bandeem, W. R., M. Havel, and I. Strange, 1965: A Radiation Climatology in the Visible and Infrared from the TIROS Meteorological Satellites, NASA TN D-2534.
- Climate and Weather of Southeastern Asia, 1942: Part 1, India, Burma, and Southern China, Vol. 5, No. 3, United States Government Printing Office, Washington.
- Colon, J. A., 1964: On Interactions Between the Southwest Monsoon Current and the Sea Surface over the Arabian Sea, Indian J. Met. and Geoph., Vol. 15, No. 2, 183-200.
- Conover, J. H., 1965: Cloud and Terrestrial Albedo Determinations from TIROS Satellite Pictures, J. Applied Met., Vol. 4, No. 3, 378-386.
- Chakravorty, D. C., and S. C. Basu, 1957: The Influence of Western Disturbances on the Weather over Northeast India in Monsoon Months, Indian J. Met. and Geoph., Vol. 8, No. 2, 261-272.
- Food and Agriculture Organization of the United Nations, 1968: The State of Food and Agriculture 1968, Rome, 147-155.
- Fritz, S. and J. S. Winston, 1962: Synoptic Use of Radiation Measurements from Satellite TIROS II, Mon. Wea. Rev., Vol. 90, No. 1, 1 - 9.
- House, F. B., 1965: The Radiation Balance of the Earth from a Satellite, Ph.D. thesis, Department of Meteorology, University of Wisconsin.
- NASA Staff Members, 1964: TIROS VII Radiation Data Catalogue and Users' Manual, National Aeronautics and Space Administration, Washington.
- Pant, P. R., 1964: Onset of Monsoon over India, Indian J. Met. and Geoph., Vol. 15, No. 3, 375-380.
- Raman, C. R. V., 1964: Monsoon Definitions, Indian J. Met. and Geoph., Vol. 15, No. 2, 235-238.
- Rao, P. K., 1966: A Study of the Onset of the Monsoon over India During 1962 Using TIROS IV Radiation Data, Indian J. Met. and Geoph., Vol. 17, No. 3, 347-356.

- Rao, P. K. and J. S. Winston, 1963: An Investigation of Some Synoptic Capabilities of Atmospheric "Window" Measurements from Satellite TIROS II, J. Applied Met., Vol. 2, No. 1, 12 - 23.
- Shamshad, K. M., 1967: Note on a Technique of Long Range Forecasting for Monsoon Rain in West Pakistan, J. Applied Met., Vol. 6, No. 1, 199-202.
- Simpson, G. C., 1921: The Southwest Monsoon, Quart. J. Roy. Met. Soc., Vol. 47, 151.
- Sparkman, B. B., 1964: Experimental Analysis of the TIROS Hemispheric Sensor, M. S. thesis, Department of Meteorology, University of Wisconsin.
- Staff Members, 1964: Weather, Monsoon Season (June - September 1963), Indian J. Met. and Geoph., Vol. 15, No. 1, 101 - 108.
- Staff Members, 1965: Weather, Monsoon Season (June - September 1964), Indian J. Met. and Geoph., Vol. 16, No. 1, 141 - 148.
- Staff Members, 1966: Weather, Monsoon Season (June - September 1965), Indian J. Met. and Geoph., Vol. 17, No. 1, 115 - 120.
- Suomi, V. E., 1958: The Radiation Balance of the Earth from a Satellite, Annals of the IGY, Vol. 1, 331-340.
- Suomi, V. E., K. J. Hanson and T. H. Vonder Haar, 1967: The Theoretical Basis for Low-Resolution Radiometer Measurements from a Satellite, Annual Report. Grant WBG-27, Department of Meteorology, University of Wisconsin, 79 - 100.
- Vonder Haar, T. H., 1968: Variations of the Earth's Radiation Budget, Ph. D. thesis, Department of Meteorology, University of Wisconsin.
- Wagner, A., 1931: Zur Aerologic des Indischen Monsuns, Ger. Beitr. z. Geoph., Bd. 30, 196.
- Wark, D. Q., G. Yamamoto and J. H. Lienesch, 1962: Methods of Estimating Infrared Flux and Surface Temperature from Meteorological Satellites, J. Atm. Sc., Vol. 19, No. 5, 369 - 384.
- Yin, M. T., 1949: A Synoptic-Aerologic Study of the Onset of the Summer Monsoon over India and Burma, J. Met., Vol. 6, No. 6, 393 - 400.

N71-11623

SIMULATION AND ANALYSIS OF MICROWAVE PROPAGATION
BETWEEN OCCULTATION SATELLITES

BY

Douglas H. Sargeant
University of Wisconsin

ABSTRACT

This report presents the results of studies designed to evaluate the feasibility of remotely probing the earth's atmosphere by measuring the refraction of microwaves propagated between occultation satellites.

Several methods are developed for accurate calculation of refraction variables, and techniques for the representation of vertical refractivity profiles are explored. The results are applied to the simulation of microwave refraction for a number of atmospheric conditions, both modeled and "real."

Finally, the inversion of microwave occultation measurements to obtain useful meteorological data is discussed.

TABLE OF CONTENTS

	Page
I. Introduction	284
A. The Genesis of This Report	284
B. The Status of the Occultation Technique	285
C. Summary of Pertinent Relations	287
II. Numerical Evaluation of Refraction Variables	289
A. Summary	289
1. Analytical Evaluation	289

2. Numerical Quadrature	289
3. Ray-tracing	289
B. Theoretical Considerations	289
1. Numerical Quadrature	289
2. Ray-tracing	291
C. Comparative Results for Exponential Profiles	294
III. Representation of Refractivity Profiles	294
A. Philosophy	294
B. Forms of Representation	296
1. Some Selected "Real" Profiles	296
2. Mathematically Simple Analytical Profiles	298
3. Generalized Analytical Forms	298
4. Application to Real Profiles	307
IV. Results of Refraction Simulation Studies	313
A. Numerical Quadrature	313
1. Introduction	313
2. Specific Results	314
B. Ray-tracing	315
1. Introduction	315
2. Accuracy Criteria and Layer Selection	318
3. Ray-tracing with Specified Minimum Ray Heights	325
4. Ray-tracing with Specified Satellite Positions	326
V. The Inversion Problem	334
A. Introduction	334
B. General Considerations	339
C. Mathematical Inversion	340
D. Statistical Inversion	340
Appendix A—Analytical Calculation of Laguerre Coefficients	342
Appendix B—Calculation of the Contribution to Laguerre Coefficients Due to the Upper Portion of the Path	343
Acknowledgments	344
References	344

1. INTRODUCTION

A. The Genesis of This Report

This study was begun early in 1967 with the stimulation of V. E. Suomi in response to a definite need. The successful Mars occultation experiments achieved using the Mariner probes had stimulated interest in possible

applications for observation of the Earth's atmosphere. A widely circulated proposal prepared by a Stanford University student group described a specific system (Stanford, 1966) and raised a large number of interesting scientific and technological questions. A climate of keen interest in novel, potentially versatile global observing systems had already been generated by the international and national proposals for the World Weather Watch and the Global Atmospheric Research Program (GARP). However, only a few groups which included Suomi and Stanford's B. B. Lusignan's appeared prepared and motivated to undertake the required feasibility studies. The required studies include not only substantial engineering systems design but also meteorological, propagation and data analysis problems.

In particular, the author decided to make careful simulation calculations as applied to the Earth's atmosphere. This required refinement of certain approximations that were applicable to the Martian atmosphere. The investigation of the occultation technique using "real" data, and in particular Southern Hemisphere data, was also proposed. The National Science Foundation approved a grant to carry out this work in Australia during the academic year 1967-68. The study was begun during the summer of 1967 with the support of an ESSA grant supervised by Suomi and was continued in Australia. A preliminary report was prepared in January, 1968, and published (Sargeant, 1968). Some of the salient results are presented in Section C, "Summary of Pertinent Relations."

The detailed study continued until the author's departure from Australia, but a final report was not completed. Fortunately, almost immediately upon return, the writer was invited to join a NASA Study Group, formed to evaluate the occultation technique, so that the results of the work were disseminated informally to interested parties. During the summer of 1969, this final report was prepared with the partial financial support of a NASA grant supervised by Suomi.

Although many of the results obtained early in the study are now dated and have served their utility as foundations for further work, some remain of current interest and others are worthy of record for future reference to avoid unnecessary duplication of effort. Already published results will not be belabored here. Thus, for a basic summary of the occultation concept, see the paper by Lusignan, et al. (1969).

B. The Status of The Occultation Technique

The early excitement and enthusiasm of the promise of an Earth-atmosphere occultation technique has waned as the detailed problems have been systematically identified and explored. To my knowledge, none of the technical problems identified appears to be insurmountable. However, the

"climate" has changed because of scientific, technological and operational developments. In particular, the specification of observational requirements and the concept of a global observing system have evolved considerably during the past few years of planning for GARP. An incomplete set of observations, inhomogeneous in space, time, nature and quality, now appears acceptable. Significant achievements in observing systems include routine sounding by infrared spectrometers on an orbiting satellite and wind determination from cloud motions observed by geostationary satellites. Further improvements are firmly scheduled.

On the other hand, the concept of a global horizontal balloon system has proven inapplicable throughout most of the troposphere and faces problems in the tropics due to probable systematic export of balloons to higher latitudes. An adequate global system for water vapor observation remains to be found. The determination of a pressure-height reference level is still in question. In fact, the whole problem of substituting, interchanging and combining the mass distribution or temperature field and the wind field still is not resolved. Different trade-offs and accuracy requirements appear to apply at different latitudes. It is in this uncertain context that we must assess a possible role for occultation measurements.

The most natural meteorological variable to consider (and the initial source of interest) is the mass distribution. In dry air, the refraction effects depend solely and quite sensitively on the density distribution. Furthermore, the microwave measurements can be obtained down to the surface even through clouds, in contrast to optical and infrared measurements. However, the appeal of these apparent advantages has been moderated by two factors. First, the simplest microwave measurements are significantly affected by water vapor in the region below cloud level, while the region above is already accessible to infrared measurements. Second, the application of statistical inversion methods to the NIMBUS III SIRS data has allowed inference of profiles below clouds to an extent exceeding original expectations, and the results should be substantially improved by the use of sensors with higher spatial resolution such as now being prepared for future satellites.

One should not immediately rule out any significant role for occultation observation of the density field, however. For one thing, the vertical resolution of occultation measurements should be far better than infrared measurements. For another, by combining suitable complementary measurements, such as attenuation near 1.35 cm, the water vapor contribution may be separable so that profiles can be recovered much lower in the atmosphere. Even climatological estimates of water vapor provide a substantial improvement. Both of these approaches have been investigated by Pomalaza (1969), and some of the results are described in Section V.

In the lower troposphere, where water vapor is abundant, refraction effects are dominated by the vapor distribution. Rather than despair of the complica-

tions thereby introduced for the deduction of density, perhaps we can deduce the water vapor profiles! Only recently has the concept of partial, complementary measurements come into vogue. Thus, even if occultation measurements of phase path and attenuation cannot define both the density and vapor profiles, perhaps the addition of IR-derived temperature profiles will allow a solution. Such possibilities still require investigation.

Another possible contribution of even a single pair of occultation satellites is the determination of a pressure-height reference level. That is, given temperature as a function of pressure above some unknown height (for example, as derived by IR sounding above cloud level), both profiles are to be determined as a function of height. A preliminary study by the Stanford group (Lusignan, 1970) has already yielded promising results.

At the present time, the meteorological-technological community apparently has set aside the microwave occultation method as inapplicable in the near future. It is not obvious to the writer that the concept should yet be completely buried. The basic philosophy of the present study has been to try to prove that the occultation method won't work, by systematically devising a series of necessary conditions. The failure to meet any of these conditions would condemn the method to death. As yet, I have been unable to kill it.

C. Summary of Pertinent Relations

The basic relations for ray analysis of propagation in a spherically stratified atmosphere were developed in detail in the preliminary report by Sargeant (1968). Some of the essential results are summarized here.

In geocentric polar coordinates (r, θ) , where θ is measured from the point of closest approach $(\hat{r}, 0)$, and with refractive index $n = n(r)$, the ray path satisfies Snell's law:

$$nr \cos \alpha = \hat{n} \hat{r} = \text{constant} \quad (1.1)$$

where α is the angle between the ray and the circles of constant n . The ray is described by the differential equation

$$r \frac{d\theta}{dr} = \cot \alpha = \frac{\hat{n} \hat{r}}{n(r)} (n^2 r^2 - \hat{n}^2 \hat{r}^2)^{-1/2}. \quad (1.2)$$

The bending of the ray due to non-uniformity of the medium, denoted by $\epsilon = \theta - \alpha$, is determined by the relation $d\epsilon = -\cot \alpha \, d \ln n$, so that

$$\epsilon(r) = -\hat{n} \hat{r} \int_{\hat{r}}^r \frac{n'(r) dr}{n(r)(n^2 r^2 - \hat{n}^2 \hat{r}^2)^{1/2}}. \quad (1.3)$$

The optical path length of a ray, denoted by L , is defined in terms of the ray

arc length s such that $dL = nds = nr \sec \alpha d\theta$. It is convenient to express L as the sum of terms

$$L(r) = \mathcal{L}(r) + \ell(r) \quad (1.4)$$

where

$$\begin{aligned} \mathcal{L}(r) &= (n^2 r^2 - \hat{n}^2 \hat{r}^2)^{1/2} + \hat{n} \hat{r} \epsilon(r) \\ &= nr(\sin \alpha + \epsilon \cos \alpha) = \hat{n} \hat{r} (\tan \alpha + \epsilon) \\ &= nr[\sin \theta(1 + \epsilon^2/2 - \epsilon^4/8 + \dots) - \cos \theta(\epsilon^3/3 - \epsilon^5/30 + \dots)] \quad (1.5) \end{aligned}$$

and

$$\ell(r) = - \int_r^R (n^2 r^2 - \hat{n}^2 \hat{r}^2)^{1/2} \frac{n'(r)}{n(r)} dr. \quad (1.6)$$

Note that ϵ , L , \mathcal{L} and ℓ are all measured from the point of closest approach and are therefore equal to half of the total values for propagation between satellites.

Using these equations, the primary refraction variables can be conveniently and accurately calculated for a specified refractive index profile, height of closest approach and terminal height. For the case of satellite terminals, the computations of ϵ and ℓ become entirely independent of satellite position (except insofar as that would determine \hat{r}) because both integrands are proportional to the gradient $n'(r)$ which becomes negligible at great heights. This allows us to isolate the effects of refractive profiles from those of satellite positions.

The calculation of \mathcal{L} (and hence L), of course, requires specification of satellite position. However, L will be on the order of a few thousand kilometers, while phase path measurement resolutions of a few centimeters are contemplated. Thus calculations must be very accurate to be maximally useful. By using equations (1.4) and (1.5), only the relatively small variable portions of L need be calculated by elaborate numerical procedures. For example, for the earth's atmosphere ϵ , ℓ and the excess of L over the geometrical "line of sight" are on the order of 10 mrad, 100 m and a few hundred meters, respectively. This greatly relaxes the requirement for numerical resolution in the calculation of ϵ and ℓ compared to that for L . Incidentally, these considerations also explain the utility of the last form of equation (1.5), which can tolerate much less accurate values of ϵ than the other forms and which does not require values of \hat{n} and \hat{r} at all.

II. NUMERICAL EVALUATION OF THE REFRACTION VARIABLES

A. Summary

Three different methods have been employed to obtain numerical values for the various refraction integrals; analytical evaluation, numerical quadrature and ray-tracing. These methods are briefly summarized and compared below.

1. Analytical Evaluation

This method has been applied to the analysis for exponential refractivity profiles. It is fully developed in the preliminary report (Sargeant, 1968). The primary advantage of the method is the ease with which the effects of variation of the parameters, both "signal" and error, may be investigated.

2. Numerical Quadrature

This method is convenient when the profile is specified analytically. It is very efficient when the profile is smooth and similar to an exponential profile.

3. Ray-tracing

When a profile is specified in tabular form, as are most empirical profiles, a ray-tracing procedure appears to be favored. However, care must be exercised to insure that a sufficiently accurate procedure is employed in view of the stringent resolution requirements.

B. Theoretical Considerations

1. Numerical Quadrature

It was shown by Sargeant (1968) that for exponential profiles the desired integrals for ϵ and l have the form $\int_0^{\infty} x^{-1/2} e^{-x} f(x) dx$ and $\int_0^{\infty} x^{1/2} e^{-x} g(x) dx$, respectively, where $f(x)$ and $g(x)$ are well-behaved functions of order unity. Although these integrals were successfully evaluated using a rather complicated analytical procedure which yielded a rapidly convergent series, the method could not be generalized readily to other profile forms. However, integrals of this form can be evaluated conveniently using a Gauss-Laguerre quadrature scheme; that is,

$$\int_0^{\infty} x^{-1/2} e^{-x} f(x) dx \doteq \sum_{i=1}^m w_i f(x_i) \quad (2.1)$$

where the x_i are the zeros of the generalized Laguerre polynomial of degree m , viz. $L_m^{(-1/2)}(x)$, and w_i are the corresponding quadrature weights. Such a formula is exact when $f(x)$ is a polynomial of degree less than $2m$ and thus is very efficient (see e. g., Hildebrand, 1956). The required zeros and weights are tabulated, e. g., in the IBM system/360 Scientific Subroutine Package, Programmer's Manual, or they may be generated by the method described in Abramowitz and Stegun (1964).

Referring to equations (1.3) and (1.6), and letting $x = (r - \hat{r})/H$ where H is a suitable scale height, we find that

$$\epsilon(r) = \hat{n} \left(\frac{r}{2H} \right)^{1/2} \int_0^X x^{-1/2} e^{-x} \frac{[-He^{-x} N'(\hat{r} + Hx)]}{[1 + N(\hat{r} + Hx)]^2} \left\{ 1 + \frac{Hx}{2\hat{r}} - \frac{\hat{r}}{Hx} \cdot \left[1 + \frac{N(\hat{r}) + N(\hat{r} + Hx)}{2} \right] \frac{[N(\hat{r}) - N(\hat{r} + Hx)]}{[1 + N(\hat{r} + Hx)]^2} \right\}^{-1/2} dx \quad (2.2)$$

$$l(r) = (2\hat{r}H)^{1/2} \int_0^X x^{-1/2} e^{-x} [-Hxe^{-x} N'(\hat{r} + Hx)] \cdot \left\{ 1 + \frac{Hx}{2\hat{r}} - \frac{\hat{r}}{Hx} \left[1 + \frac{N(\hat{r}) + N(\hat{r} + Hx)}{2} \right] \frac{[N(\hat{r}) - N(\hat{r} + Hx)]}{[1 + N(\hat{r} + Hx)]^2} \right\} dx \quad (2.3)$$

where $X = (r - \hat{r})/H$. Sargeant (1968) has shown that the upper limit may be extended to infinity without appreciable error for a satellite terminal. Thus we expect the quadrature procedure to be efficient provided that

$$e^{x} N'(\hat{r} + Hx) \quad \text{or} \quad e^{\frac{r-\hat{r}}{H}} N'(r)$$

is well approximated by a polynomial, which is true if $N(r)$ is similar to a decreasing exponential of scale height H .

For the special case of $N(r) = \hat{N}e^{-\frac{(r-\hat{r})}{H}} = \hat{N}e^{-x}$, these formulas become

$$\epsilon = \hat{n} \hat{N} \left(\frac{\hat{r}}{2H} \right)^{1/2} \int_0^{\infty} x^{-1/2} e^{-x} (1 + \hat{N}e^{-x})^{-2} \left\{ 1 + \frac{Hx}{2\hat{r}} - \frac{\hat{N}\hat{r}}{H} \left[\frac{1 + \frac{1}{2}\hat{N}(1+e^{-x})}{(1+\hat{N}e^{-x})^2} \right] \left(\frac{1-e^{-x}}{x} \right) \right\}^{-1/2} dx \quad (2.4)$$

$$l = \hat{N}(2\hat{r} + H)^{1/2} \int_0^{\infty} x^{-1/2} e^{-x} x \left\{ 1 + \frac{Hx}{2\hat{r}} - \frac{\hat{N}\hat{r}}{H} \left[\frac{1 + \frac{1}{2}\hat{N}(1+e^{-x})}{(1 + \hat{N}e^{-x})^2} \right] \left(\frac{1-e^{-x}}{x} \right) \right\}^{1/2} dx \quad (2.5)$$

so that the computation is particularly simple. (In this case, accuracy is enhanced by letting $F(x) = f(x) - 1$ and computing the ϵ -integral as

$$I = \frac{\pi}{2} + \int_0^{\infty} x^{-1/2} e^{-x} F(x) dx, \quad \text{where } f(x) \text{ is as in equation (2.1). A similar procedure can be used for } l.)$$

This method is also effective for the much more general case of a profile which can be represented in the following form:

$$N(h) = e^{-h/H} \sum_{k=0}^p a_k \left(\frac{h}{H} \right)^k \quad \text{where } h = r - \hat{r}$$

i. e., an exponential times a polynomial of degree p . For reasons discussed in Section III, B. 3, it is convenient to use

$$N(h) = e^{-h/H} \sum_{k=0}^p b_k L_k \left(\frac{h}{H} \right) \quad (2.6)$$

where L_k is the Laguerre polynomial of degree k . The L_k can be evaluated using the recursion relation

$$\begin{aligned} L_{k+1}(x) &= (2k+1-x)L_k(x) - k^2 L_{k-1}(x) \\ L_0(x) &= 1, \quad L_1(x) = 1 - x. \end{aligned} \quad (2.7)$$

The derivative can also be evaluated recursively, by noting that

$$-He^x N'(h) = e^{-h/H} \left[\sum_{k=0}^p b_k \left(\frac{h}{H} \right) - \sum_{k=1}^p b_k L_k' \left(\frac{h}{H} \right) \right] \quad (2.8)$$

and

$$xL_k'(x) = kL_k(x) - k^2 L_{k-1}(x). \quad (2.9)$$

See Section IV, A for discussion of results obtained using the quadrature method.

2. Ray-tracing Calculations

The author noted earlier that special attention is required to

insure a routine having sufficient accuracy. A number of different programs were developed. The final version was later noted to be similar to the method published by Thayer (1967). The method makes use of the fact that the refraction variables may be integrated exactly through a layer having a profile form which is power-law in r . This latter form is negligibly different from a constant-gradient layer, but is easy to apply exactly.

From equations (1.1) through (1.6), we find that if we introduce the symbol

$$\Omega \equiv -\frac{r}{n} \frac{dn}{dr} = -\frac{d \ln n}{d \ln r} \quad (2.10)$$

then:

$$d\alpha = (1 - \Omega)d\theta \quad (2.11)$$

$$d\epsilon = \Omega d\theta = [\Omega/(1-\Omega)]d\alpha \quad (2.12)$$

$$\begin{aligned} dL &= \hat{n} \hat{r} \sec^2 \alpha d\theta = [\hat{n} \hat{r} / (1-\Omega)] d(\tan \alpha) \\ &= \hat{n} \hat{r} [d(\tan \alpha) + d\epsilon] + \hat{n} \hat{r} \tan^2 \alpha d\epsilon \end{aligned} \quad (2.13)$$

$$\text{and} \quad d\ell = \hat{n} \hat{r} \tan^2 \alpha d\epsilon = [\hat{n} \hat{r} \Omega / (1-\Omega)] d(\tan \alpha - \alpha). \quad (2.14)$$

Thus for any layer through which Ω can be considered constant, the functions θ , ϵ , ℓ and L can all be integrated in closed form. Only the values of α at the boundaries of the layer are required, and these can be determined using Snell's Law (equation (1.1)) and the given profile.

The assumption that $\Omega = \text{constant}$ implies that for any layer

$$\frac{n_2}{n_1} = \left(\frac{r_2}{r_1}\right)^{-\Omega}. \quad (2.15)$$

It follows that constant- Ω layers having thickness $r_2 - r_1$ small compared to the earth radius are virtually identical to linear layers. If the layer thickness is chosen suitably small compared to the local scale height, a profile can be adequately represented in this form. Once suitably thin layers have been selected, the corresponding Ω -values and related layer constants can be calculated once and for all for a given profile. Methods for selecting and interpolating suitable profile values are discussed in section IV. B. 2. If the k^{th} layer extends from r_k to r_{k+1} ,

$$\Omega_k = \frac{\ln[n(r_k)/n(r_{k+1})]}{\ln(r_{k+1}/r_k)}. \quad (2.16)$$

For computational purposes, this can be written

$$\Omega_k = \frac{x}{y} \left[\frac{1 - \frac{1}{2}x + \frac{1}{3}x^2 - \frac{1}{4}x^3 + \dots}{1 - \frac{1}{2}y + \frac{1}{3}y^2 - \frac{1}{4}y^3 + \dots} \right] \quad (2.17)$$

where

$$x = \frac{N(h_k) - N(h_{k+1})}{1 + N(h_{k+1})} \quad \text{and} \quad y = \frac{h_{k+1} - h_k}{r_0 + h_k}.$$

In this notation, the refraction variables can be calculated

$$e = \sum_k \left(\frac{\Omega_k}{1 - \Omega_k} \right) (\alpha_{k+1} - \alpha_k) \equiv \sum_k A_k \Delta \alpha_k; \quad (2.18)$$

$$l = \hat{n} \hat{r} \sum_k \frac{\Omega_k}{1 - \Omega_k} [(\tan \alpha_{k+1} - \tan \alpha_k) - (\alpha_{k+1} - \alpha_k)] \equiv \hat{n} \hat{r} \sum_k A_k \Delta \lambda_k; \quad (2.19)$$

$$L = \hat{n} \hat{r} \sum_k \frac{1}{1 - \Omega_k} [\tan \alpha_{k+1} - \tan \alpha_k] = \hat{n} \hat{r} \sum_k B_k \Delta \Lambda_k. \quad (2.20)$$

The quantities $\Delta \alpha_k$, $\Delta \lambda_k$ and $\Delta \Lambda_k$ can be calculated as follows:

From equation (1.1)

$$\cos \alpha = \frac{\hat{n} \hat{r}}{n r} = \left(1 + \frac{\hat{N} - N}{n}\right) \left(1 + \frac{\hat{h} - h}{r}\right);$$

define

$$e(h) \equiv \frac{h - \hat{h}}{r_0 + h} - \frac{\hat{N} - N}{1 + N} + \left(\frac{h - \hat{h}}{r_0 + h}\right) \left(\frac{\hat{N} - N}{1 + N}\right) \quad (2.21)$$

so that

$$c_k = \cos \alpha_k = 1 - e(h_k). \quad (2.22)$$

But

$$e = 1 - \cos \alpha = 2 \sin^2 \frac{\alpha}{2}$$

so that

$$\Delta \alpha_k = 2 \left[\sin^{-1} \left(\frac{e_{k+1}}{2} \right)^{1/2} - \sin^{-1} \left(\frac{e_k}{2} \right)^{1/2} \right]. \quad (2.23)$$

Also,

$$\Delta \Lambda_k = \frac{\sin \Delta \alpha_k}{c_k c_{k+1}} \quad (2.24)$$

and

$$\begin{aligned}\Delta\lambda_k &= \frac{\sin \Delta\alpha_k}{c_k c_{k+1}} - \Delta\alpha_k \\ &= \frac{\Delta\alpha_k}{c_k c_{k+1}} \left[(1 - c_k c_{k+1}) - \frac{1}{3!} (\Delta\alpha_k)^2 + \frac{1}{5!} (\Delta\alpha_k)^3 + \dots \right].\end{aligned}\quad (2.25)$$

See section IV. B for the discussion of results obtained using this ray-tracing method.

C. Comparative Results for Exponential Profile.

Table 1 shows some comparative results for rays passing at selected minimum heights (\hat{h}) in an atmosphere having exponential refractivity $N = N_0 e^{-h/H}$, with $N_0 = 250$ N units and $H = 8$ km. These values might roughly approximate a dry atmosphere.

The Gaussian quadrature method is expected to be virtually exact in this case. The analytical method appears to provide the accuracy predicted in the preliminary report, viz. 0.1% or better. The two ray-tracing programs differ principally in their method of layer selection. Both results are included in order to illustrate how important it is to use an adequate procedure. Details of the selection procedure and related accuracy criteria are discussed in section IV. B. 2.

Table 1 certainly justifies confidence that all of the various methods and programs are logically correct.

III. REPRESENTATION OF REFRACTIVITY PROFILES

A. Philosophy

Although an exponential refractivity-profile model allows us to make reasonable estimates for the magnitudes and expected variations of the refraction effects on a radio signal propagated between satellites, such a parameterization is only a crude representation. The estimates of surface pressure and scale height of the Martian atmosphere provided by preliminary analysis of occultation data from the Mariner IV probe constituted a significant increase in knowledge, but such estimates for the earth's atmosphere would be only mildly interesting at best. If "occultation" data are to prove useful for meteorological analysis, much more detailed information must be extracted. This obviously implies that more complete parameterizations are required both for simulation and for inversion studies.

Table 1.

Comparative Refraction Computations Using Alternate Methods for an Exponential Refractivity Profile
($N_0 = 250$ N units, $H = 8$ km)

Minimum Ray Height (h) in km	Bending Angle (ϵ) in mrad				Excess Retardation lengthening (ℓ) in meters			
	Approximate Analytical	Gaussian Quadrature	Ray Trace A	Ray Trace B	Approximate Analytical	Gaussian Quadrature	Ray Trace A	Ray Trace B
0	9.676	9.6753	9.6749	9.6755	66.466	66.480	66.484	66.376*
1	8.442	8.4406	8.4403	8.4406	59.099	59.129	59.132	59.025
2	7.376	7.3752	7.3750	7.3753	52.511	52.537	52.540	52.435
4	5.654	5.6527	5.6525	5.6527	41.353	41.374	41.376	41.275
8	3.357	3.3567	3.3566	3.3564	25.464	25.476	25.477	25.377
17	1.211	1.2111	1.2110	1.2105	9.506	9.510	9.510	9.416

*See Section IV. B. 2 for explanation of differences

It is somewhat difficult to separate entirely considerations appropriate for simulation from those for inversion. The most realistic propagation simulations will of course result from the most detailed tabulations of empirical profiles. We shall see that such simulations essentially require a ray-tracing approach. On the other hand, simple analytical representations of profiles provide definite advantages for studying the effects of variations of profile parameters. Perhaps more significantly, presently conceivable experiments with earth-orbiting satellites can provide only a limited (and small) number of radio parameters for inversion purposes, roughly corresponding to observing only a few discrete heights in the atmosphere.

For these reasons, it is necessary to investigate the problem of finding simple representations for vertical profiles, if only to attack the inversion problem. The discovery described earlier that a simple, low-order numerical quadrature calculation provided a fast and accurate determination of the refraction variables for a quasi-exponential refractivity profile naturally led to a consideration of parameterizations which utilize this technique. Simulation based on quadrature is quite convenient and accurate for a sizable class of profiles, but an excessively large number of parameters is required if adequate representation and simulation is attempted for arbitrary empirical profiles.

It now appears that a more statistical approach—for example, using empirical orthogonal functions—is superior for representation, whereas straightforward ray-tracing using a tabulated profile is convenient for simulating refraction effects. Nonetheless, we shall present some of the results of the parameterization studies as they provide insight into the inversion problem.

B. Forms for Representation

Profiles may be represented either analytically or in tabular form. Analytical profile formulations may be selected primarily for mathematical simplicity, for fidelity or for some combination of these considerations. Tabular profiles are usually required to describe "real" observed profiles accurately.

1. Some Selected "Real" Profiles

It is useful to consider a basic set of "typical" profiles for use in both representation and simulation problems. In keeping with the basic philosophy of this research, the initial set was chosen to minimize the problem of simulation, interpretation and inversion of "occultation" data. In particular, it was desired to employ profiles which might apply over a large horizontal area so that the assumption of spherical stratification would be reasonably valid, and for which the contribution due to water vapor would not be overwhelming. On the other hand, it was desired to include detailed profile features such as inversions and strong-gradient layers so that a realistic assessment of potential difficulties could be made. On these bases, a number

of composite profiles were chosen which are typical of various synoptic situations as observed at Campbell Island near 53°S latitude in the Antarctic Ocean. These profiles were originally derived for another purpose by Martin (1968) from three months of daily winter soundings in 1966. The original temperature and dew point profiles are shown in Fig. 1. Note that all profiles have been normalized to 1000 mb. Typical surface pressures are also given.

For each of the profile types, microwave refractivity profiles were calculated using the familiar Smith-Weintraub formula (see Bean and Dutton, 1966)

$$N = 77.6 \frac{P}{T} + 3.73 (10^5) \frac{e}{T^2} \equiv N_{\text{dry}} + N_{\text{wet}} \quad (3.1)$$

A formula essentially equivalent to the Magnus formula was used to represent the saturation vapor pressure function. The formula used reproduced the values tabulated in the Smithsonian Tables to nearly four significant digits over the range of temperatures encountered. Heights were calculated by upward integration of the hypsometric equation using mean virtual temperature for each layer.

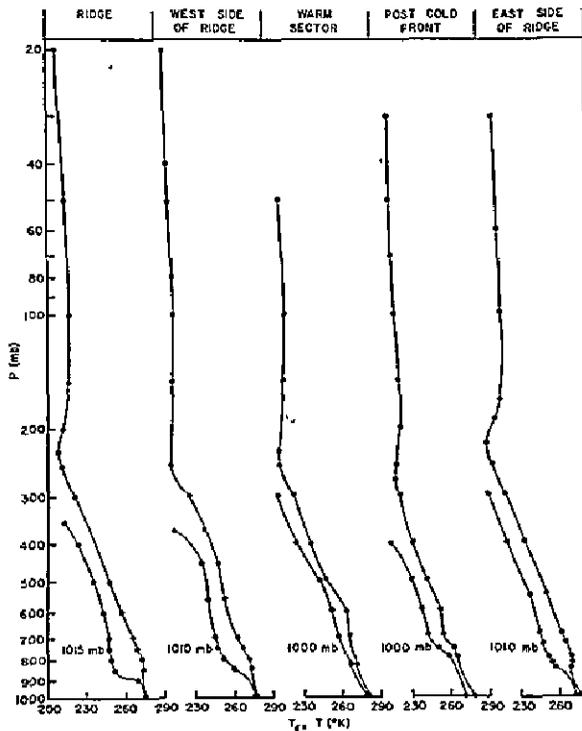


Fig. 1. Some typical "real" temperature and dewpoint temperature profiles. Composite soundings, normalized to 1000 mb but with typical surface pressure given, for Campbell Island (53°S) in the Antarctic Ocean.

6

The "dry" and "wet" terms were also calculated separately to assess the water vapor contribution. The results are shown in Figs. 2a through 2e. Some of the same data are plotted in Fig. 2f on a semi-log graph. This shows the quasi-exponential nature of refractivity profiles.

When absolute heights must be compared between these profiles, an adjustment for surface pressure is necessary. In this report, this was accomplished by shifting the entire profile upward by 10 m/mb and extrapolating a new value at the surface.

2. Mathematically Simple Analytical Profiles

A number of simple profile representations that have proved useful in past studies are summarized in the Radiometeorology monograph of Bean and Dutton (1966). These include

- (1) Linear or "effective earth radius" Model: $dN/dz = \text{constant}$
- (2) Exponential model: $N(z) = N_0 e^{-z/H}$
- (3) Bi-exponential model: $N(z) = N_1 e^{-z/H_1} + N_2 e^{-z/H_2}$ (interpretation: separate exponential "dry" and "wet" terms)
- (4) Statistically parameterized models: Exponential, piecewise linear and exponential, etc.

See the cited monograph for further discussion, typical values of the parameters, and literature references. Some examples are shown in Fig. 3. These forms were used in some of the trial calculations described below.

3. Generalized Analytical Forms

For various purposes, especially for the mathematical inversion of radio refraction measurements, it is desirable to employ a more general analytical representation. The problem is to represent more "realistic" profiles but still retain a limited number of parameters. It is common to employ power series in such problems, but as refractivity profiles are generally quasi-exponential functions of height, many terms would be required to account for that fact alone. The following representation, motivated partially by the numerical quadrature method described above, employs an exponentially weighted polynomial:

$$N(z) \simeq \mathcal{N}(z) = e^{-z/H} \sum_{k=0}^m a_k z^k \equiv e^{-z/H} p_m(z) \quad (3.2)$$

This form assumes that the weighted profile $e^{z/H} N(z)$ can be efficiently represented by a polynomial for some choice of scale height H . The profiles

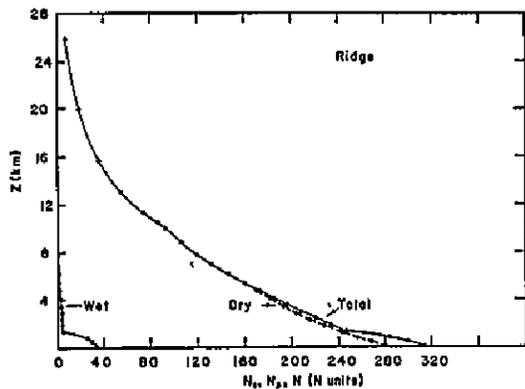


Fig. 2a

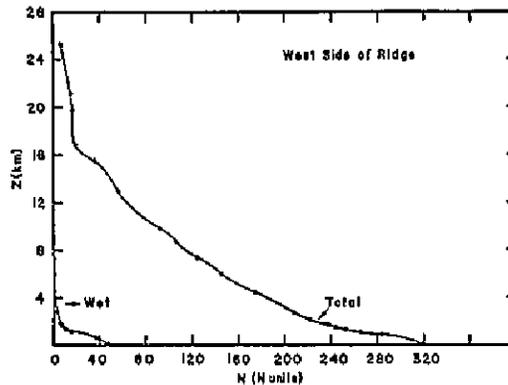


Fig. 2b

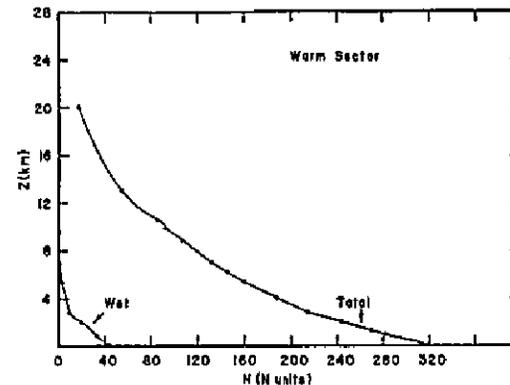


Fig. 2c

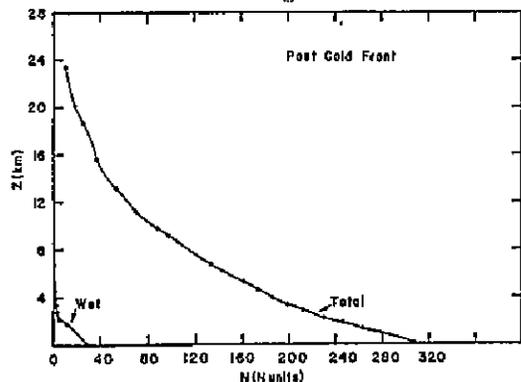


Fig. 2d

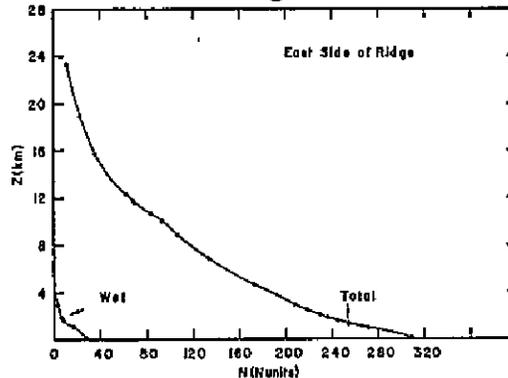


Fig. 2e

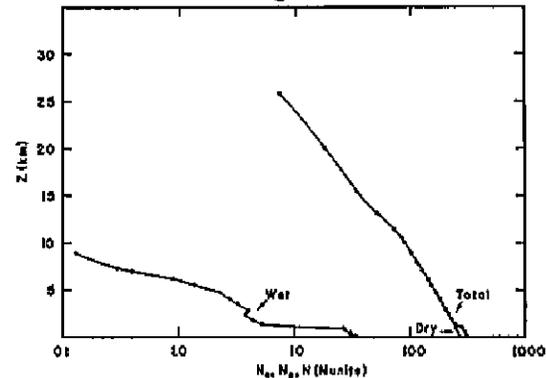


Fig. 2f

Fig. 2a-e. Calculated total refractivity and wet component of refractivity for "real" profiles shown in Fig. 1.

2f. Semi-log plot of refractivity vs. height for ridge profile showing quasi-exponential nature of profile.

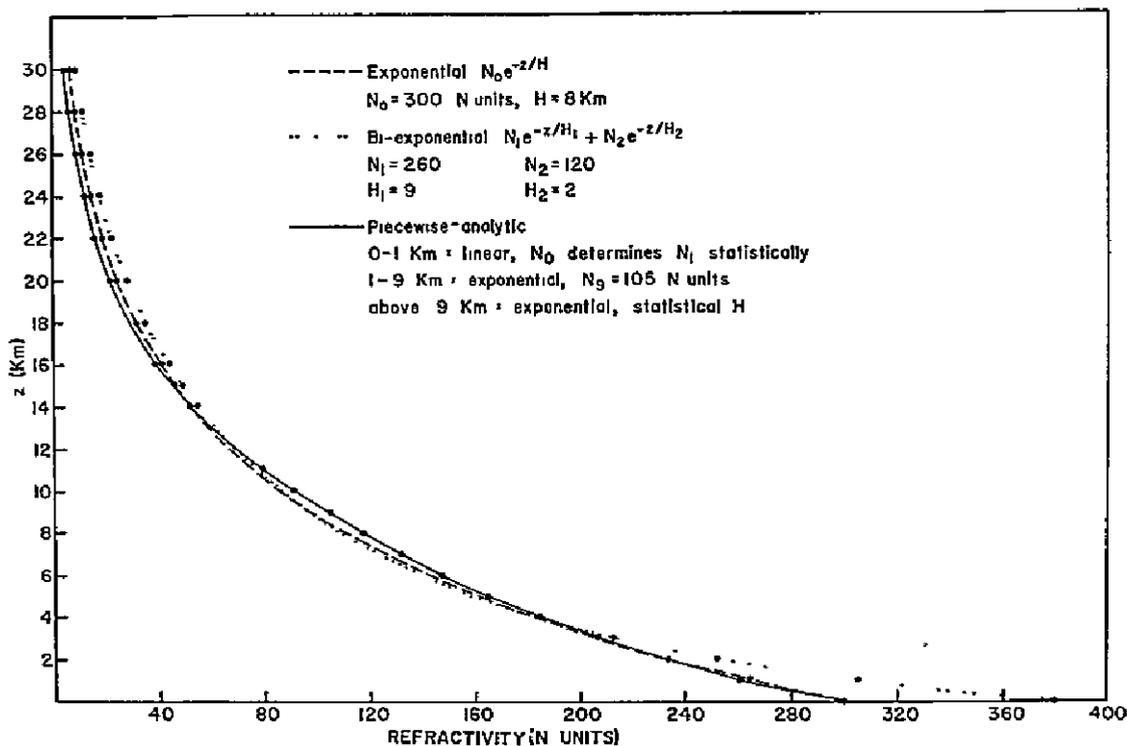


Fig. 3. Examples of some analytic refractivity profiles.

shown in Fig. 4a-e correspond to those of Fig. 2a-e with exponential weighting. Note that detailed profile features such as inversions are retained and even emphasized after weighting, and that it is difficult to compensate the general lapse in both the troposphere and stratosphere by use of a single scale height.

The form of equation (3.2) requires the specification of $m + 2$ parameters (H and the a_k). The question arises: how can these parameters be efficiently determined for a given profile? In most of the following, it is assumed that a suitable value for H can be determined independently, or perhaps by an iterative optimization procedure.

a) Fitting specified "anchor points." Perhaps the simplest procedure is to determine the a_k such that $p_m(z)$ agrees with $e^{z/H}N(z)$ at $m + 1$ specified points. This method was investigated numerically, but the results are summarized only briefly here. The general conclusion is that successful representation of quasi-exponential, analytically specified profiles is quite possible using this technique, but is, as expected, rather sensitive to the particular set of "anchor points" selected.

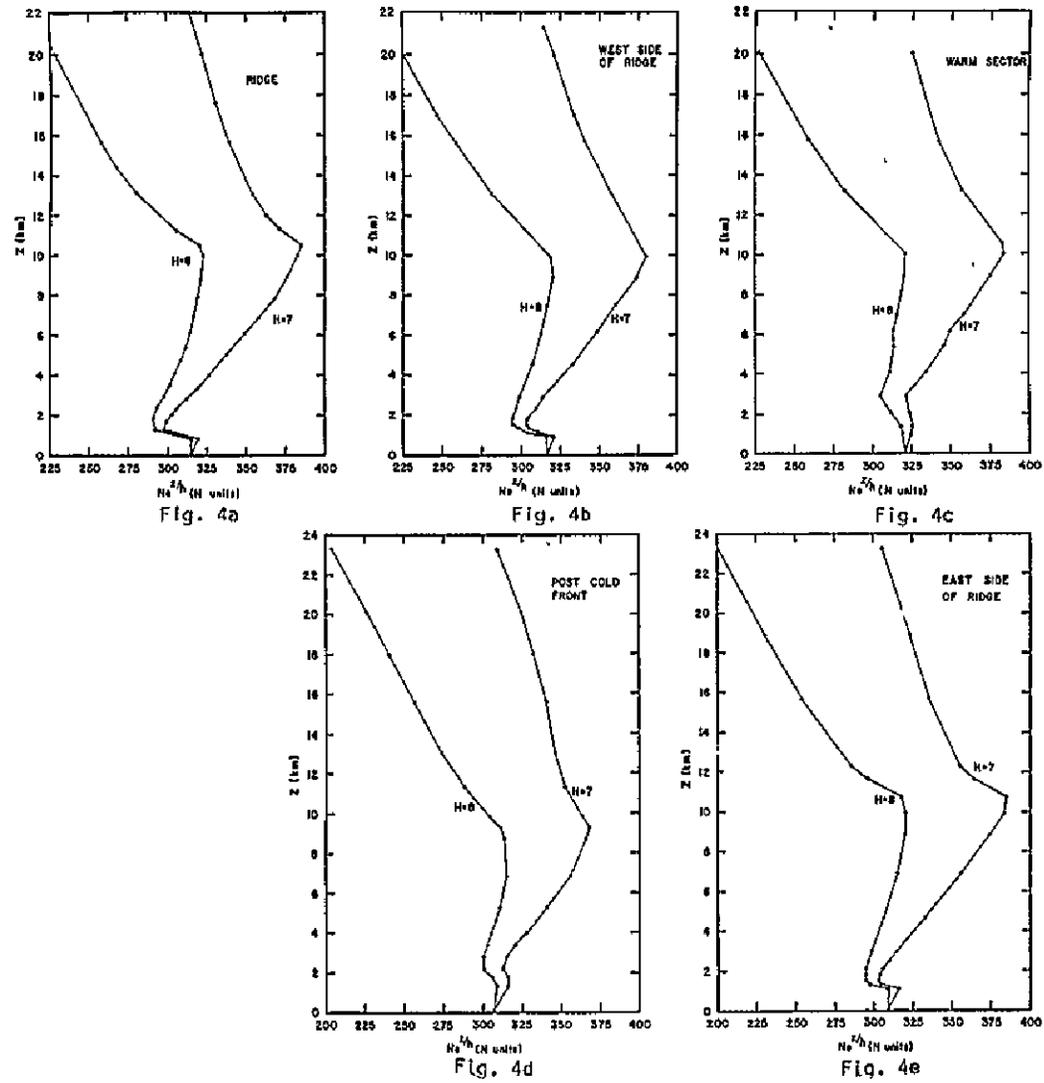


Fig. 4a-e. Some examples of exponentially weighted refractivity profiles.

Calculations were made using a selection of exponential, bi-exponential and piecewise-analytic profile models, and for various choices of the "compensating" scale height H and sets of anchor points. The results were qualitatively insensitive to variations in H of a few kilometers. The exponential profiles could be well-fitted over the range 0 - 30 km by almost any reasonable choice of a small number (5 or 7) of anchor points, while the piecewise-analytic profiles could be fitted reasonably only by a judicious choice of anchor points. Even the smooth bi-exponential profiles yielded "wild" results (e. g. negative N in some intervals) for particularly unfortunate choices of anchor points. A low-order fit to an empirical profile would undoubtedly produce even more problems.

It was, therefore, concluded that an objective procedure for point selection or profile fitting was highly desirable, if not essential.

b) Optimization for $z \in (0, \infty)$: Laguerre Profiles. We can choose the polynomial coefficients in order to minimize the weighted-square error over the entire profile:

$$\xi = \int_0^{\infty} e^{-z/H} [N(z) - \tilde{N}(z)]^2 dz.$$

It is then convenient to write

$$N(z) = e^{-z/H} \sum_{k=0}^m b_k L_k(z/H), \quad (3.3)$$

where $L_k(\chi)$ denotes the Laguerre polynomial of degree k which satisfies the normalization $\int_0^{\infty} e^{-\chi} L_k^2(\chi) d\chi = (k!)^2$, and also the recursion relations given in Section II, B. 1. In this case, ξ is minimized by choosing

$$b_k = \frac{1}{H(k!)^2} \int_0^{\infty} N(z) L_k(z/H) dz = \frac{1}{(k!)^2} \int_0^{\infty} N(H\chi) L_k(\chi) d\chi. \quad (3.4)$$

The minimum value of ξ afforded by the above polynomial is

$$\xi_{\min} = H \left[\int_0^{\infty} e^{\chi} N^2(H\chi) d\chi - \sum_{k=0}^m (k! b_k)^2 \right].$$

This parameter can be employed in criteria for selecting values of H and m .

The coefficients b_k must somehow be calculated for a given profile. In the case of simple analytical profiles such as exponential and bi-exponential, the integrals may be evaluated analytically. It is also possible to optimize the

choice of H analytically. These calculations are presented in Appendix A, and the results were used as a check on the more general method described below.

For a more complicated profile, the integral for b_k must be calculated numerically. Since the basic assumption of the entire representation procedure is that $e^{z/H}N(z) = e^{\chi N(H\chi)}$ is well-approximated by a polynomial, it is natural to consider a Gauss-Laguerre quadrature scheme similar to that already used in Section II. B. 1. In this case, we calculate

$$(k!)^2 b_k = \int_0^\infty e^{-\chi} [e^{\chi N(H\chi)} L_k(\chi)] d\chi \doteq \sum_{i=1}^M W_i e^{\chi_i} L_k(\chi_i) N(H\chi_i) \quad (3.5)$$

where the χ_i are the zeros of the polynomial $L_M(\chi)$ described in Section II. B. 1, and W_i are the corresponding quadrature weights.

In order to perform the calculations in (3.5), it is only necessary to know the values of refractivity $N(z)$ at the M discrete heights $H\chi_i$. This is encouraging from the standpoint of the inversion problem, provided that M can be kept small. It turns out that when M is small, the gross characteristics of the profile are extracted, but M cannot remain small if fine-scale features of the profile must be accurately described. Note that this requirement on M is different from that which determines the adequate value for m in equation (3.3). In that case, assuming that the b_k are known exactly, the quality of the "fit" depends on the number of coefficients used ($m+1$). The latter consideration is significant when employing the numerical quadrature method of refraction simulation as described in Section II. B. 1.

In any case, the M special values of $N(H\chi)$ suffice only to calculate coefficients b_k for $k < M$, and values b_k obtained are accurate only to the extent that $e^{\chi N(H\chi)}$ really is well-approximated by a polynomial of degree $2M - 1 - k$ or less. That is, if $e^{\chi N(H\chi)}$ really were a polynomial of degree m then $e^{\chi N(H\chi)} L_k(\chi)$ would be a polynomial of degree $m+k$; and the highest degree coefficient b_m would require a scheme accurate for polynomials of degree $2m$, viz. a value M such that $2M - 1 > 2m$.

Computer programs were written to perform efficiently the calculations required in (3.5) and (3.3), making use of the recursion relations of Section II. B. 1. Again, exponential, bi-exponential and piecewise-analytic functions were considered first. The effects of choice of compensation scale height H , the number of "input" specified points on the profile, and the number of coefficients in the representation were all studied with emphasis on recovery of the profile between 0 - 30 km.

The "stability" problems of fitting specified "anchor points" are thus completely removed. Purely exponential profiles are again easy to recover. A guess of H within 1 km yields accuracy of 4 significant digits in the recovered

profile for 3 or 4 point quadrature formulas. Even bad guesses for H require only a few more points, and better guesses require only two. (Clearly, if we knew a priori that the profile was exponential, only two points would be required to determine the two parameters.)

For other than exponential profiles, some of the considerations hinted at above become important, particularly if both high accuracy and low-order representations are sought. First, a good choice of H is important, especially if the effective scale height of the profiles varies considerably throughout the range of interest. Second, the number of "observed" points (M) must be large enough to accurately determine a particular coefficient b_k , and this depends on k . Third, the number of coefficients ($m + 1$) must be large enough to describe variations of the scale desired. Finally, some of the high-order coefficients b_k which can be calculated using a given M may actually degrade the recovered profile.

It is convenient to illustrate these points using the bi-exponential profile shown in Fig. 3 because analytical solutions are available for comparison from Appendix A. Figure 5 shows the deviations between the original profile and those obtained using scale heights of 7 and 8 km, and also those obtained using the theoretically optimum scale heights derived in Appendix A. The coefficients obtained from the low-order quadratures are used for $H = 7$ and 8, but only exact coefficients are used for the optimum cases. The importance of scale height selection is obvious from the graphs as is the superior description of steep gradients offered by more coefficients. Table 2 compares the first few coefficients obtained by various order quadrature formulas with the exact coefficient values. Note that the highest degree coefficients obtainable with a given order quadrature may be seriously in error, perhaps by a factor of two or three.

These basic difficulties are compounded when the above technique is applied to "real" profiles. If the profile has pronounced local features, even low-degree coefficients may be inaccurate, even though high-order quadrature formulas are used. The source of the difficulty is perhaps better understood by considering the nature of the Laguerre polynomials. A given quadrature formula "samples" a profile at heights determined by the zeros of one of these polynomials. Figure 6 shows the lower "sample heights" for a few quadrature orders, assuming a scale height of 7 km. Note that even high order formulas make only a few samples of the region of interest, and thus may miss local "wiggles."

Since most of the interesting profile variations occur in a limited height range, coefficients may be calculated more accurately by using a finer-scale integration procedure in a finite lower height interval. This has been done using both trapezoidal and Gauss-Legendre quadrature, and the results are presented in Section III. B. 4 on applications. Of course, it is necessary in

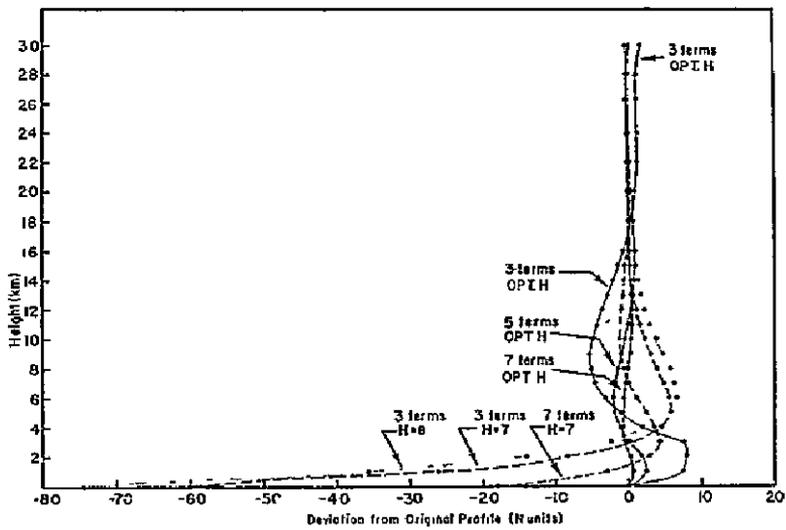


Fig. 5. Quality of Laguerre fit to bi-exponential profile with an extreme lower gradient. See Fig. 3.

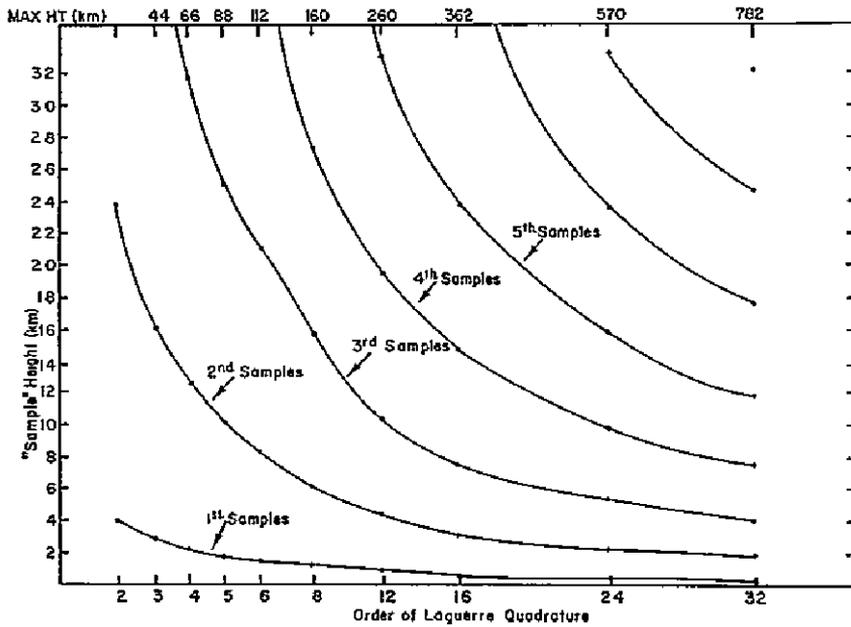


Fig. 6. Distribution of heights sampled by various order Laguerre quadrature formulas for $H = 7$ km.

Table 2. Laguerre coefficients obtained by Gauss-Laguerre quadrature of a bi-exponential profile, using a representation scale height of $H = 8$. (See Fig. 3 for the profile.)

M*	b_0	b_1	b_2	b_3	b_4	b_5	b_6
2	310.16	-28.909					
3	317.05	-22.284	5.326				
4	320.23	-18.101	7.770	0.723			
5	321.59	-15.936	9.207	1.292	.1603		
6	322.14	-14.899	9.983	1.635	.2649	2.394 (-2)**	
7	322.36	-14.426	10.375	1.824	.3273	3.941 (-2)	3.039 (-3)
8	322.45	-14.217	10.564	1.922	.3620	4.858 (-2)	4.953 (-3)
9	322.48	-14.127	10.652	1.970	.3804	5.371 (-2)	6.083 (-3)
10	322.49	-14.089	10.692	1.994	.3897	5.645 (-2)	6.717 (-3)
EXACT	322.50	-14.062	10.723	2.014	.3985	5.925 (-2)	7.417 (-3)

*M = number of points in quadrature formula

** (-2) denotes $\times 10^{-2}$

this case to separately account for the contribution from the upper part of the range. A procedure is developed in Appendix B.

On the other hand, Fig. 6 also tells us that even high-degree representations with accurate Laguerre coefficients will have trouble describing profile details in the lower region of interest. The Laguerre optimization procedure simply requires too many of its degrees of freedom to accomplish the fit all the way to infinity. This naturally leads us to consider fitting over a finite height interval, and modelling the upper portion separately.

c) Optimization for $z \in (0, z_m)$: Legendre Profiles.

Separate modelling of the upper atmosphere appears justified because simple models seem adequate there and because the upper portion of a ray path will not contribute much to the occultation measurements anyway. However, forcing a large number of "wiggles" into a finite height interval will not necessarily yield more satisfactory results for propagation simulation. Gradients of refractivity are most important, not absolute values, and gradients can be seriously distorted by a "wandering" fit. In any case, we can minimize the mean-square error between the exponentially-weighted profile and a polynomial over a finite interval:

$$\begin{aligned} \xi &= \int_0^{z_m} [e^{z/H} N(z) - e^{z/H} \eta(z)]^2 dz \\ &= z^* \int_{-1}^1 [e^{z^*/H} e^{\frac{z^* \chi}{H}} N(z^* + z^* \chi) - p_m(\chi)]^2 d\chi, \text{ where } z^* = z_m/2. \end{aligned}$$

It is convenient to write

$$N(z) = e^{\frac{-(z-z^*)}{H}} \sum_{k=0}^m c_k P_k(z/z^* - 1) \quad (3.6)$$

where $P_k(\chi)$ is the Legendre polynomial of degree k . Then the desired coefficients c_k satisfy

$$c_k = \frac{2k+1}{2} \int_{-1}^1 e^{\frac{z^*\chi}{H}} N(z^* + z^*\chi) P_k(\chi) d\chi. \quad (3.7)$$

As in the case of the Laguerre representation (Section III. B. 3. b), in analogy with part b), the integrals 3.7 for the coefficients can be calculated by any suitable numerical method. It is convenient to use Gauss-Legendre quadrature, so that

$$c_k = \frac{2k+1}{2} \sum_{i=1}^M W_i e^{\frac{z^*\chi_i}{H}} N(z^* + z^*\chi_i) P_k(\chi_i) \quad (3.8)$$

where the χ_i and W_i are the zeros and weights for an M -point Gauss-Legendre formula. Again, it may not be desirable to retain all of the calculable coefficients.

These polynomials will, of course, have all of their "samples" and "wiggles" within the specified height interval. Some results are presented below in Section III. B. 4. b.

d) Chebyshev Optimization, and so on. It should also be possible to select the polynomial p_m in order to minimize the maximum error $|e^{z/H}N(z) - p_m|$ over a given interval. This procedure has not been investigated. "Wandering fit" problems mentioned earlier can almost certainly be expected. Also, as mentioned earlier, it now appears more satisfactory to use statistical expansion functions rather than systematically testing classical orthogonal polynomials for fortuitous resemblance to actual profiles.

4. Application to Real Profiles

In the preceding section, an exponentially-weighted polynomial form was proposed for generalized representation of refractivity profiles, and some methods for determining "optimum" coefficients were described. The methods were illustrated and tested using analytically-specified profiles. In this section, the application to a "real" profile is examined. The results are illustrated using the "Ridge" profile shown in Fig. 2a, and shown exponentially-weighted in Fig. 4a. Note especially that even the weighted form displays a marked, high-gradient inversion in the lowest two kilometers and also the tropopause inversion near 10 kilometers. These features will, of course, constitute a demanding test of the representation capability. The profile above

the last tabulated value of 25.8 kilometers is assumed exponential with a scale height of 7 kilometers.

- a) Laguerre Fit. The following problems are recalled from Section III, B. 3. b:
- (i) How should the Laguerre coefficient b_k be calculated?
 - (ii) For a given method of calculation, how many of the coefficients will be sufficiently accurate to contribute usefully to the representation?
 - (iii) How many coefficients are required to describe the details of the profile to a desired accuracy?

The answers to these questions depend on the intended application. Relatively crude and simple schemes suffice to produce useful "smoothed" representations of the profile which still reveal the pertinent details. However, if very accurate representation is required to perform simulated refraction calculations, or to recover meteorological variables to high accuracy, then much effort is required.

Laguerre coefficients were calculated using the following methods: a) Laguerre quadrature of various orders (using values linearly interpolated within the exponentially-weighted table), b) Trapezoidal quadrature of various spacings interpolated within the table, supplemented by an analytical "correction" for the upper part of the profile using the method of Appendix B, and c) Legendre quadrature of various orders within the table, plus analytical "correction." The results for the first ten coefficients are summarized in Tables 3, 4 and 5.

Table 3 shows that even high-order Laguerre quadrature formulas do not converge upon stable values for the coefficients. The reason for this can be seen by considering the distribution of quadrature "sampling heights" shown in Fig. 6 in relation to the profile in Fig. 4a. Only for high-order formulas are the details of the lower inversion sampled adequately. Fortunately, even the approximate coefficients obtained by low-order formulas yield a satisfactory low-order fit. Accurate coefficients are required only for detailed fits.

To obtain more accurate coefficients, integration schemes must be used which consider the details of the profile. Table 4 shows the results obtained using Legendre and trapezoidal quadrature over the tabulated profile. Also included is the "correction" for the upper part of the profiles. Table 5 summarizes the final results obtained by the various methods.

Once a set of Laguerre coefficients has been obtained, accurate or not, a complete profile is specified. In order to display in detail the accuracy of fit, the difference between the actual and a few example Laguerre profiles plotted for the lowest 28 kilometers in Fig. 7a. Recall that the Laguerre profiles themselves are smooth, so that the excursions in Fig. 7a are due to the rapid local variations in the actual profile. Also note that the absolute error is plotted, and that N decreases from near 300 to a few N units over this range.

Table 3. Laguerre Coefficients Obtained by Gauss-Laguerre Quadrature of Ridge Profile, Using a Representation Scale Height of $H = 7$. (See Figs. 2a and 4a for the profile.)

M^*	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
2	324.7	7.4	-3.9							
3	320.1	-8.0	-3.9							
4	325.3	-14.4	-10.3	-2.0						
5	333.7	-13.5	-13.6	-3.9	-0.51					
6	331.4	-10.2	-10.6	-3.2	-0.55	-5.3 (-2)**				
7	329.0	-8.3	-8.8	-2.6	-0.46	-5.4 (-2)	-4.0 (-3)			
8	329.6	-5.2	-7.2	-2.2	-0.35	-3.8 (-2)	-2.8 (-3)	-1.4 (-4)		
9	330.9	-3.0	-6.3	-1.8	-0.27	-1.8 (-2)	4.7 (-3)	2.3 (-4)	2.0 (-5)	
10	333.1	-3.0	-6.1	-1.7	-0.20	-3.0 (-3)	3.2 (-3)	6.3 (-4)	6.7 (-5)	4.1 (-6)
12	339.4	-5.3	-8.0	-2.1	-0.21	1.2 (-2)	7.9 (-3)	1.5 (-3)	1.9 (-4)	1.8 (-5)
16	333.8	-3.1	-6.5	-1.9	-0.24	-6.6 (-3)	4.0 (-3)	9.8 (-4)	1.4 (-4)	1.6 (-5)
24	333.7	-6.4	-8.2	-2.2	-0.28	-3.6 (-3)	5.6 (-3)	1.3 (-3)	1.8 (-4)	2.0 (-5)
32	331.2	-6.2	-7.7	-2.1	-0.27	-7.7 (-3)	4.0 (-3)	9.6 (-4)	1.4 (-4)	1.4 (-5)

*M = number of points in quadrature formula

** (-2) denotes $\times 10^{-2}$

Table 4. Laguerre Coefficients for Ridge Profile Obtained by Various Methods. Representation Scale Height $H = 7$.

M	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
Obtained Using M-Point Legendre-Gauss Quadrature										
4	322.1	16.9	-22.18	-4.97						
8	327.1	21.9	-19.44	-4.03	-4.2 (-2)*	9.53 (-2)	1.79 (-2)	1.73 (-3)		
12	324.3	21.2	-19.52	-4.09	-7.9 (-2)	8.38 (-2)	1.56 (-2)	1.37 (-3)	4.5 (-5)	-4.0 (-6)
16	325.3	22.1	-19.11	-3.98	-5.4 (-2)	8.83 (-2)	1.63 (-2)	1.46 (-3)	5.4 (-5)	-3.0 (-6)
24	325.4	21.8	-19.30	-4.03	-6.1 (-2)	8.78 (-2)	1.64 (-2)	1.48 (-3)	5.8 (-5)	-2.6 (-6)
32	325.2	21.9	-19.26	-4.02	-6.2 (-2)	8.73 (-2)	1.62 (-2)	1.46 (-3)	5.5 (-5)	-2.9 (-6)
Obtained Using Trapezoidal Quadrature										
No. of Points	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
22 (Original Data)	326.96	22.48	-19.8	-4.32	-1.1 (-1)	9.1 (-2)	1.92 (-2)	2.13 (-3)	1.5 (-4)	6.0 (-6)
50	325.34	22.13	-19.1	-3.94	-3.5 (-2)	9.4 (-2)	1.75 (-2)	1.67 (-3)	8.5 (-5)	7.9 (-7)
100	325.20	21.89	-19.2	-4.01	-5.7 (-2)	8.9 (-2)	1.65 (-2)	1.51 (-3)	6.2 (-5)	-2.0 (-6)
200	325.21	21.86	-19.3	-4.03	-6.2 (-2)	8.7 (-2)	1.63 (-2)	1.47 (-3)	5.6 (-5)	-2.7 (-6)
Correction for Upper Part of Profile										
	7.48	-27.57	11.6	1.94	-0.2	-0.9	-0.12	-0.46	-0.9	-0.2

* (-2) denotes $\times 10^{-2}$

Table 5. Summary of Laguerre Coefficients for Ridge Profile Obtained by Various Methods. Representation Scale Height $H = 7$ km.

k	32-point	32-point	200-point
	Laguerre quadrature	Legendre quadrature (with correction)	Trapezoidal quadrature (with correction)
	b_k	b_k	b_k
0	331.18	332.67	332.69
1	-6.21	-5.69	-5.71
2	-7.70	-7.64	-7.65
3	-2.092	-2.088	-2.089
4	-0.269	-0.266	-0.265
5	-7.7 (-3)*	-6.0 (-3)	-5.9 (-3)
6	3.95 (-3)	4.39 (-3)	4.44 (-3)
7	9.57 (-4)	9.98 (-4)	1.01 (-3)
8	1.35 (-4)	1.45 (-4)	1.47 (-4)
9	1.43 (-5)	1.54 (-5)	1.56 (-5)

* (-3) denotes $\times 10^{-3}$

b) Legendre Fit. Legendre coefficients c_k were calculated by using Legendre-Gauss quadrature formulas of various orders and the tabulated profile over a range of 0 - 20 kilometers. For each quadrature order (M), half of the coefficients obtainable (M/2) were used to construct Legendre profiles. Some of the results are shown in Fig. 7b. The general character of low-order fits is surprisingly similar to that obtained using Laguerre profiles. As expected, the maximum deviations of the higher order Legendre profiles are somewhat smaller, but the occurrence of "wandering fit" across the true profile is more noticeable. The profiles are also clearly worthless outside the range of derivation.

c) Conclusions. A direct comparison of 10-term Laguerre and Legendre representations of the Ridge profile is shown in Fig. 8. Note that percentage deviations are given this time. It is concluded from the above studies that the proposed generalized representation provides a satisfactory method for describing smooth profiles and also for smoothing irregular profiles, but that it is not adequate to accurately describe detailed features of irregular profiles. In particular, accurate simulation of radio refraction will not be possible without using a very large number of coefficients. This fact will be demonstrated in Section IV.

It appears that an exponential profile may be simply too coarse a "first guess" to use as a base state, and that polynomials may not "wiggle" in proper

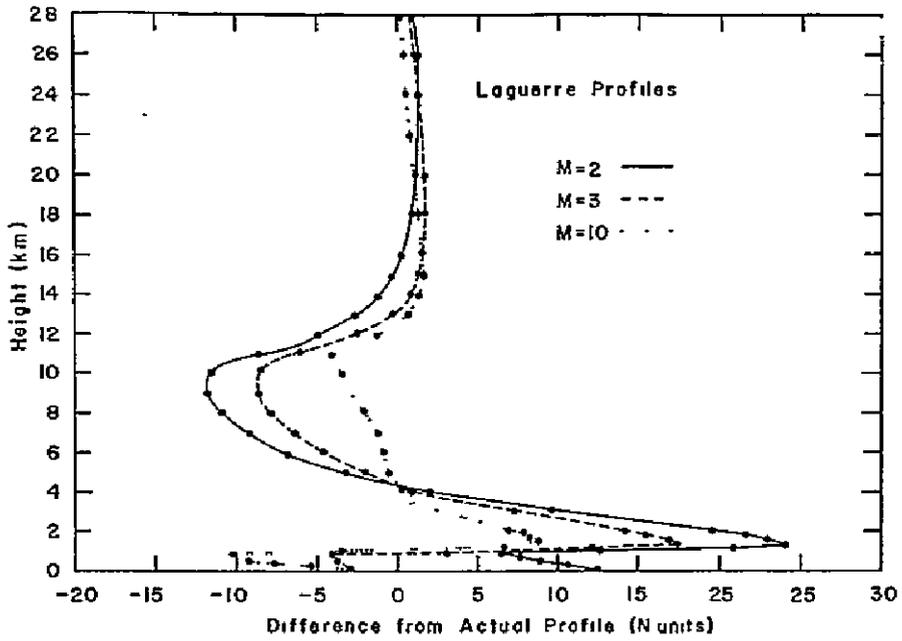


Fig. 7a

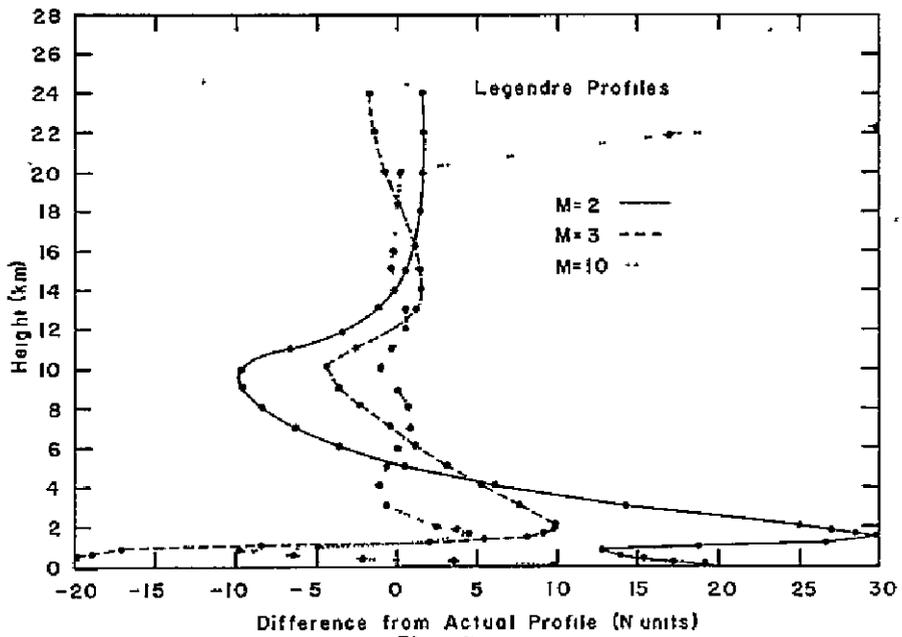


Fig. 7b

Fig. 7a and b. Quality of Laguerre (7a) and Legendre (7b) fits to the ridge profile for varying quadrature orders.

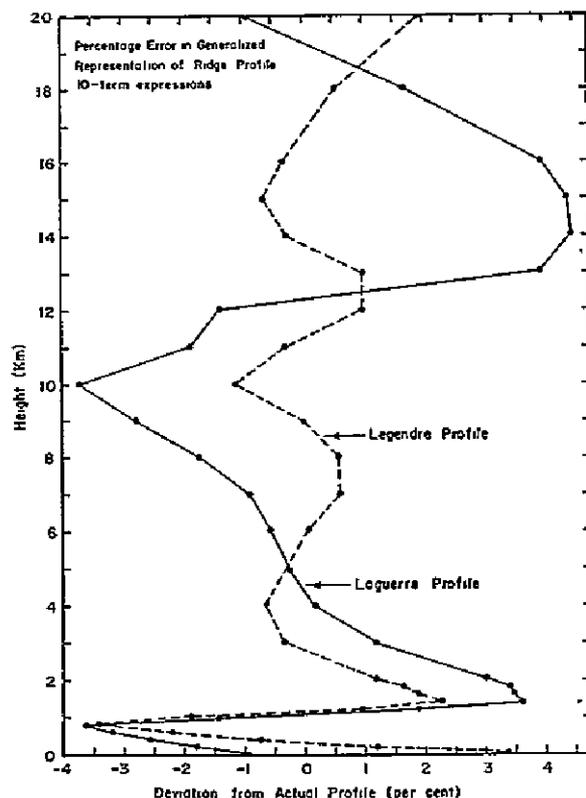


Fig. 8. Deviation of Laguerre and Legendre representations of the ridge profile. (See Fig. 2.)

places. It is suggested that the statistical mean profile relevant to a given location and time may provide a suitable base state. In this case, the appropriate expansion functions might be "empirical orthogonal functions," viz., the eigenvectors of the covariance matrix formed from a suitable collection of observed profiles. A preliminary study using six months of arctic data suggests that the first few eigenvectors, often only two, are sufficient for the description of individual profiles, including commonly occurring "details."

IV. RESULTS OF REFRACTION SIMULATION STUDIES

A. Numerical Quadrature

1. Introduction

The details of the method are given in Section II. B. 1. It was noted there that this method is highly efficient for quasi-exponential profiles. In fact, for true exponential profiles, the refraction variables ϵ and l are calculated to 3 significant digits by a 2-point formula and to 5 or 6 significant digits by a 4- or 5-point formula. This is one reason why a substantial effort was invested both in generalizing the refraction calculation and also in devel-

oping the profile representation. Unfortunately, the high efficiency is not retained for "detailed" profiles.

It is revealing to trace the evolution of the quadrature program. The first version computed ϵ and l from equations (2.4) and (2.5). The next version used equations (2.2) and (2.3), with both N and dN/dz specified analytically. The next allowed $N(z)$ to be specified by a table of Laguerre coefficients b_k , and all required values and derivatives were computed using suitable recursion relations. All versions handled exponential profiles with high efficiency. Quasi-exponential (but smooth) profiles such as the bi-exponential required higher order formulas, but were treated successfully. Profiles specified by a reasonably small number of b 's, such as the 16th order representation of "Ridge" outlined in Section III. B. 4, were also suitable.

Unfortunately, many interesting profiles can neither be specified analytically nor adequately described by a reasonable number of b 's. (Obviously, a number of b 's approaching the number of original data points is not very satisfactory.) In a final effort to extend the applicability of the method, interpolation and finite difference techniques were applied directly to the tabulated profile data. In the case of "Ridge," convergence of the ϵ and l integrals to stable values could not be obtained using quadrature formulas of 32 points or less.

2. Specific Results

The generalized quadrature method requires the specification of a suitable scale height H , just as the generalized profile representations of Section III do. As before, the efficiency of the procedure depends on the choice of H . Figure 9 shows the results of quadrature to calculate ϵ and l for a ray just grazing the earth and for the analytically-specified bi-exponential profile discussed earlier. Note that the order of quadrature formula required to obtain accurate answers is reduced by a suitable choice of H . Additional calculations with higher order formulas (not shown) establish convergence to 5 significant digits with quadrature formulas using between 12 and 16 points.

The program using Laguerre coefficients was tested using analytically derived coefficients for a true exponential of scale height 8 kilometers, but a representation H of 7 kilometers (see Appendix A). For any order of profile up to 9, the quadratures for ϵ and l converge to 5 digits by using a 5-point formula or greater. (They also converge to the correct values for the exponential profile when 5 or more coefficients are used in the profile representation.)

A further test was made using a 16th order Laguerre representation of the "Ridge" profile, with the coefficients obtained by a 32-point Laguerre formula as described in Section III. The ϵ and l quadratures converge using a 24-point formula. Figures 10a and b show values of ϵ and l versus minimum ray height for exponential, bi-exponential and Laguerre-specified "Ridge" profiles.

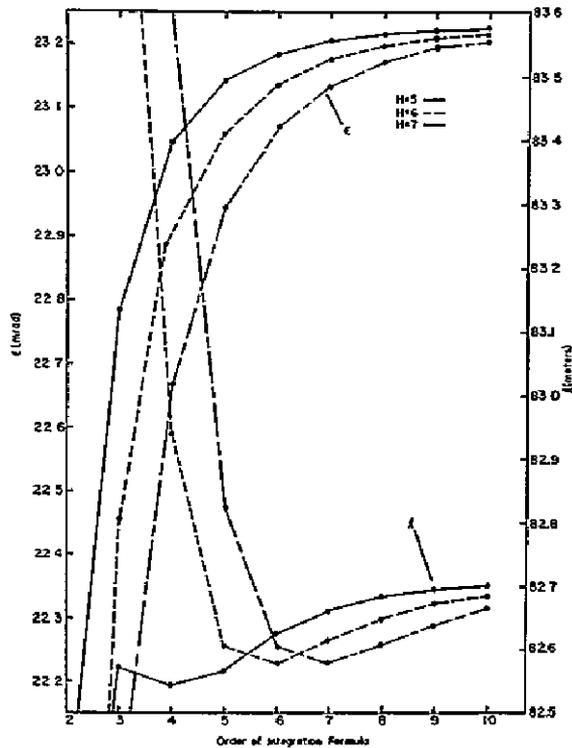


Fig. 9. Convergence properties of numerical quadrature calculation of the refraction variables, ϵ and l . Severe bi-exponential profile (Fig. 3), minimum ray height of zero km.

Table 6 gives some sample results obtained using two entirely different schemes for interpolating and differencing to estimate N and dN/dz directly from the tabulated values for "Ridge." Such results are the main reason why the quadrature method was abandoned in favor of ray-tracing for refraction simulation using real profiles.

b. Ray-tracing

1. Introduction

A ray which passes between two points (e.g., occultation satellites) is affected by the entire intervening medium, as exhibited in the integral expressions for the various refraction variables. Only the "total" path-integrated answers are available at the satellites, and the previously discussed quadrature method provides only those answers. Only those refractivity values required to define the integrals adequately for numerical quadrature are used, and intermediate terms in the summations are not retained.

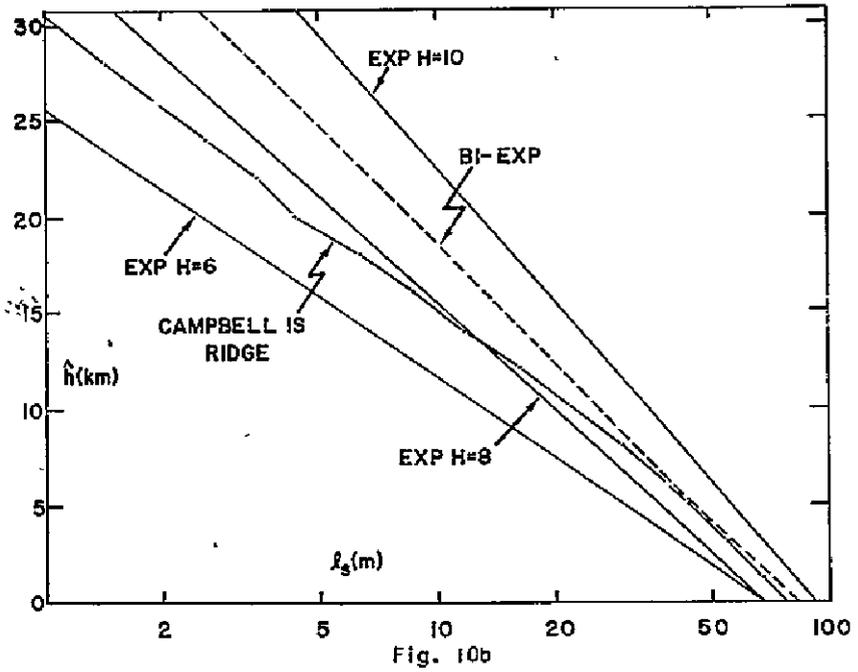
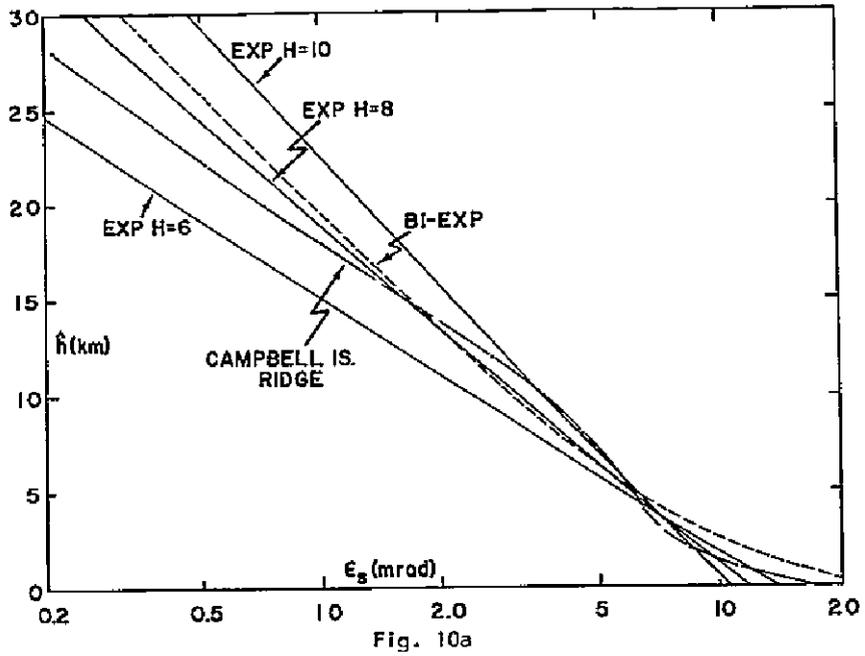


Fig. 10 a and b. Refraction variables, ϵ_s and l_s , calculated by numerical quadrature for various profiles. Campbell Island Ridge profile represented by 16th order Laguerre profile.

Table 6. Lack of convergence of refraction calculations for "Ridge" profile, specified by tabular values:

~A: linear interpolation of exponentially weighted profile
 B: exponential interpolation
 M: order of quadrature formula

M	ϵ (mrad)		l (m)		
	A	B	A	B	
4	25.26	25.37	94.81	94.94	Minimum ray height
8	11.93	11.95	76.34	76.46	Zero Kilometers
12	11.29	11.29	73.26	73.29	
16	11.20	11.20	76.67	76.57	
24	11.71	11.71	74.03	73.97	
32	16.62	16.53	88.44	88.27	

Ray-tracing not only provides the "final" answers but also determines relevant propagation variables at intermediate points along the ray-path. Thus if a detailed specification of the profile is required to do the quadratures, it may well be superior to ray-trace and thereby obtain the added intermediate information, which can, for example, be used to calculate weighting functions.

The required high accuracy of the "final" answers has been discussed in Section I. This places special requirements on a ray-tracing procedure. First, the profiles to be traced must be specified with great precision, even though this implies fictitiously high accuracy in the original measurements. Second, the ray-tracing equations employed must be very accurate, as even minor approximations may prove disastrous for differential refraction calculations. Finally, the numerical calculations themselves must be formulated so that round-off errors are unimportant.

The ray-tracing equations discussed in Section II. B. 2 are exact when applied to "power-law" layers—which are practically indistinguishable from linear-N layers. Only refractivity values at the layer boundaries are required. Thus we need only insure that each layer in the ray-tracing table is adequately represented by a linear N-profile. Other, more complicated criteria were also studied, but these are more relevant to analytically specified profiles for which the deviation from a "power-law" actually is known. Some examples are briefly discussed below.

Another problem peculiar to the occultation analysis is the requirement that rays intersect predetermined points. As discussed in Section I, this problem can be isolated somewhat by calculating ϵ and l which are affected only by the lower part of the ray-path. However, it is not possible to select, a priori, the initial ray angle (or equivalently, the minimum ray height) which will put the ray (or possibly rays) through a predetermined point. An efficient iterative procedure was developed to allow calculation of refraction variables when both the satellite height and separation are specified.

2. Accuracy Criteria and Layer Selection

a) Criteria based on the ray-trace results. The various ray-tracing versions developed were tested initially using exponential profiles to allow comparison with the other methods. Some results are presented in Table 1. Once a suitable, accurate set of ray-tracing equations is devised, the main problem is the selection of layers thin enough to insure accuracy but thick enough to keep the total number of layers manageable. Results obtained using two methods are given in Table 1.

The integrand in the ϵ formula (equation (1.3)) is singular at the minimum ray height, so that ϵ accumulates very rapidly in the lowest part of a ray. This is one reason why it is more convenient to use one of the angular variables θ or α as the ray parameter to be incremented. In one of the earliest programs, denoted RAYTRACE B in Table 1, layers were selected by incrementing θ so as to produce ϵ increments of 0.1 mrad per step in the lower part of the ray, but limiting θ steps to 1 mrad maximum. The trace was terminated when ϵ failed to increase by more than 10^{-4} mrad at any step. Between 100 and 200 steps were required depending on the minimum ray heights, and the traces terminated when the rays reached a height of about 61 km. Referring to Table 1 and using the quadrature results as a standard, we note that RAYTRACE B results for ϵ are comparably accurate except for large h , and that the l values are noticeably in error. The errors occur because ϵ continues to increase slowly at great heights when h is large and also because an appreciable fraction of l is produced at great heights in all cases. (See the weighting functions shown in Fig. 11.) Thus the traces terminate too soon.

The method denoted RAYTRACE A in Table 1 was devised to accept tabulated as well as analytically-specified profiles. The results cited in Table 1 were obtained using exponential profile values specified only at the 22 tabulated heights for the "Ridge" profile and at the assumed satellite height of 300 km. Layers were added by "halving" in the original table and interpolating exponentially until additional layers failed to alter the cumulative values of ϵ or l by a factor exceeding 10^{-5} . About 100 steps were required in each case, with several steps between 60 and 300 km. The tabulated results confirm our expectation of agreement with the quadrature method to within a few units in the fifth significant digit.

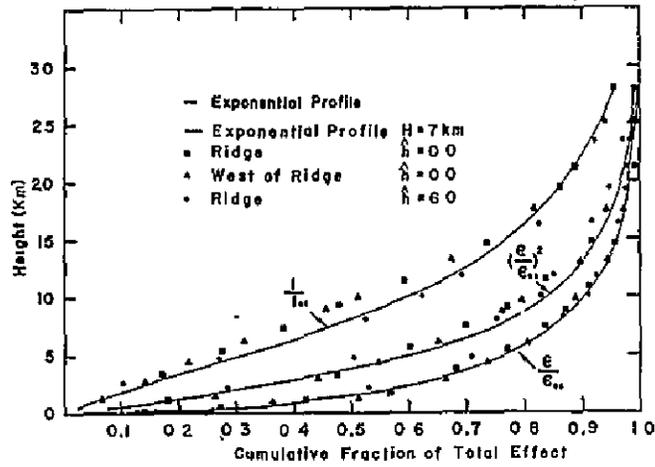


Fig. 11a

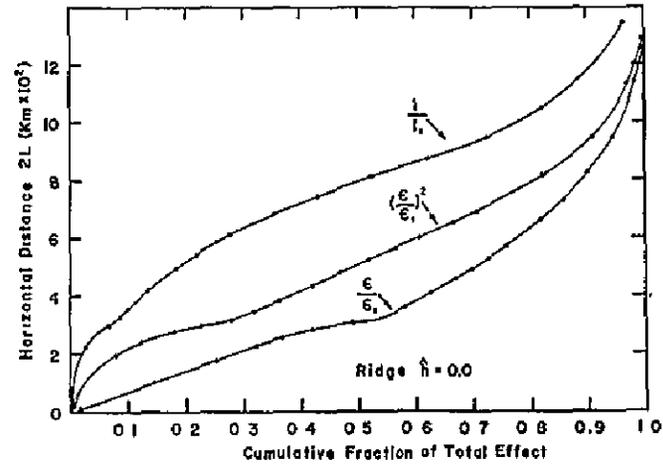


Fig. 11b

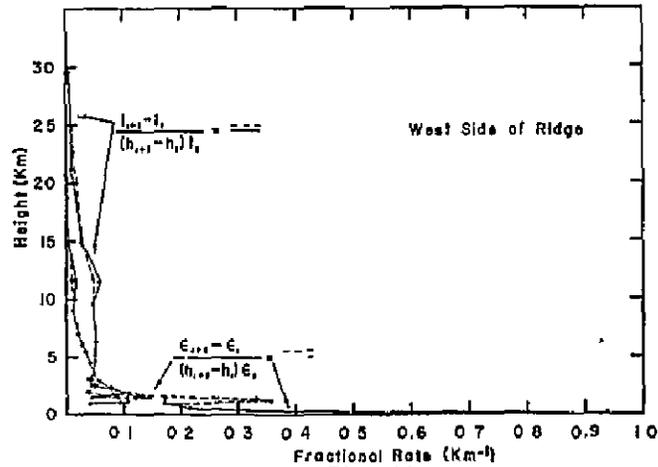


Fig. 11c

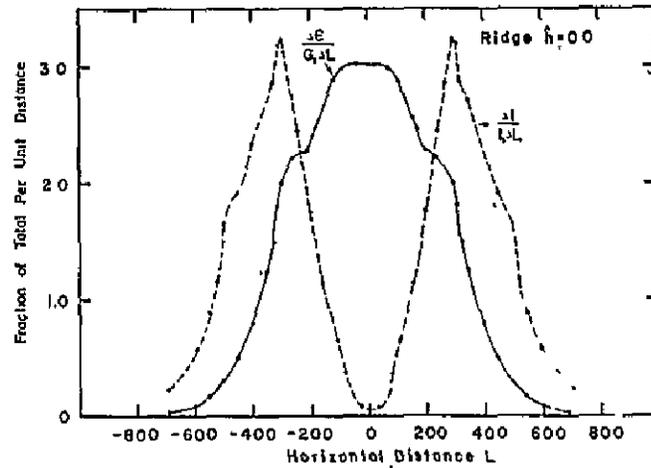


Fig. 11d

Fig. 11a - d. Examples of distribution of refraction effects along selected rays

The same method (RAYTRACE A) was also applied to "real" profiles. Table 7 compares the results obtained using values of 10^{-5} and 10^{-4} for the "convergence criterion" applied to both ϵ and l when selecting each layer as described above. Note that approximately twice as many layers are required to obtain the additional significant digit. Also note that even though ϵ and l are precisely calculated using either criterion, the values of L can differ by many meters, and this variation exceeds the resolution expected in the radio measurements. If L were calculated using equations (1.4) and (1.5), rather than by ray-tracing, a variation of only about 20 m or less in L would be expected for 1 mrad variation in ϵ ! This assumes, however, that the "line of sight distance" or satellite separation is known exactly. When this distance (or equivalently the central angle θ) is determined by ray-tracing, a difference of only 10^{-3} mrad in θ (or ϵ) produces a difference of about 6 m in $r \sin \theta$ and thus in L . This consideration is explored in later sections.

b) Criteria based on the refractivity profile. The method described above requires many repeated tests during any trace and necessitates selection of a different set of layers for each minimum ray height, even within the same refractivity profile. As mentioned earlier, it is possible to select layers once and for all, declare them to be exactly "power-law" between the specified points, and then have the equations developed in Section II. B. 2 apply exactly.

If the profile is originally derived from rather widely-spaced, vertical samples of P , T and e , then perhaps the best estimate for the refractivity N at an untabulated height will be obtained by exponential interpolation. Thus, if we assume

$$N(z) = N_1 \exp\left(-\frac{z - z_1}{H}\right) \text{ when } z \in [z_1, z_2], \text{ then}$$

$$N_E(z) = N_1 \left(\frac{N_2}{N_1}\right)^x \text{ where } x = \frac{z - z_1}{z_2 - z_1}, \quad (4.1)$$

This should be compared with linear interpolation which requires

$$N_L(z) = N_1 \left[1 - \left(\frac{z - z_1}{z_2 - z_1}\right) \left(1 - \frac{N_2}{N_1}\right) \right]. \quad (4.2)$$

It is, therefore, sufficient to construct a table for ray-tracing by merely adding enough points to the original table using equation (4.1) so that the linear assumption (4.2) in the refined table differs negligibly from (4.1). The resulting "linear" layers can be declared exactly "power-law" with as much justification as any alternative assumption.

If we let $p = (N_1 - N_2)/N_1$, then the maximum difference between equation (4.1) and equation (4.2) for a layer is given by

Table 7. Comparison using RAYTRACE A Method with Different Criteria. Ridge Profile

A1 Criterion 10^{-5}

A2 Criterion 10^{-4}

Minimum Ray Height \hat{h} (km)	No. of Layers		Bending Angle (ϵ) in mrad		Excess Retardation Lengthening (l) in meters		Total Path Length (L) in km	
	A1	A2	A1	A1 - A2	A1	A1 - A2	A1	A2
0	101	48	13.0683	.00094	78.445	-.021	2054.133	.027
1	121	54	14.9478	.00029	68.629	-.020	2063.762	.022
2	113	51	8.1223	.00049	61.754	-.013	2017.960	.014
4	108	50	6.6333	.00048	49.162	-.011	2002.868	.013
8	105	51	4.4052	.00027	29.722	-.0048	1977.019	.0089
16	104	50	1.3684	.00009	8.719	-.0015	1933.042	.0028

$$\left(\frac{N_L - N_E}{N_1}\right)_{\max} = 1 - \left[\frac{p}{-\ln(1-p)}\right] \left\{1 + \ln\left[\frac{-\ln(1-p)}{p}\right]\right\}. \quad (4.3)$$

When $p \ll 1$, this simplifies to

$$\left(\frac{N_L - N_E}{N_1}\right)_{\max} \cong \frac{1}{8} \left(\frac{N_1 - N_2}{N_1}\right)^2 \quad (4.4)$$

Thus, it is a relatively simple matter to specify a criterion for a "thin enough" layer. The results presented below were obtained by requiring $(N_1 - N_2)/N_1 < .0$ for each layer so that linear interpolations would be within about 0.03% of exponential interpolations within the refined tables, which corresponds to about 0.1 N-unit near the ground. This criterion resulted in tables of around 100 values for the earlier presented Campbell Island data.

c) Comparative Results. Table 8 provides a comparison between results obtained using the criteria of RAYTRACE A described in Section IV, B. 2, a and those obtained using the criterion described in Section IV, B. 2, b, the method henceforth designated as RAYTRACE C. Several facts should be noted. First, method A requires essentially the same total number of (different) layers no matter what the minimum ray height is, whereas method C requires only that a fraction of the permanently specified layers is actually traversed by the ray. Second, although we have insured that the two profiles actually traced are essentially the same, the differences in the final results, especially for L, are substantially greater than the anticipated resolution of the radio measurements. Third, the differences between the methods are at least somewhat systematic, with larger differences in ϵ and L for low-passing rays, and nearly constant differences in l for all rays in a given profile.

These results can be understood by considering the evolution of the differences along the path of a single ray. Some examples are given in Table 9 for two contrasting cases. As noted earlier, the bending (ϵ) occurs most rapidly near the minimum ray height and also in any region where the gradient is large. In these regions, method A requires many layers to obtain a stable value, and the computation is very sensitive to the small differences between the numerous exponentially interpolated layers and the thicker piecewise "power-law" layers used in method C. (Note that neither interpolated profile is more justified by the sparse original data, but that C is merely more convenient.) The total path length L amplifies the differences in ϵ by a sort of "optical lever" effect. Thus for a given ray the differences in ϵ and L are accumulated almost entirely within the first layer of the source profile, and the total differences for a ray are larger when the gradients in that first layer are larger (see Table 9 and Fig. 2).

Table 8. Ray-trace Comparisons with Different Layer Criteria

\hat{h} (km)	No. of Layers		ϵ (mrad)		l (m)		L (km)	
	A*	C**	C	A - C	C	A - C	C	A - C
0	103	80	13.208	.048	80.626	-.170	2054.909	.300
1	112	77	11.590	.020	71.340	-.169	2042.173	.128
2	110	74	9.908	.026	63.068	-.169	2029.011	.162
4	110	68	6.993	.007	49.788	-.169	2005.073	.042
8	104	57	4.393	.004	30.044	-.172	1976.949	.020
16	104	33	1.382	-.0004	9.083	-.182	1933.123	-.003

Data above for warm sector profile

	A	C	Value	Difference	Value	Difference	Value	Difference
	0	100	92	12.139	.043	78.852	-.104	2048.360
1	108	90	10.875	.0025	69.925	-.099	2037.799	.014
2	109	87	9.628	-.025	61.879	-.097	2027.367	-.159
4	111	79	6.773	.011	48.945	-.100	2003.736	.072
8	107	69	4.510	.007	29.117	-.100	1977.724	-.045
16	103	43	1.375	-.001	8.889	-.105	1933.077	-.006

Data above for post cold front profile

* Obtained using the 10^{-5} criteria of RAYTRACE A described in Section IV. B. 2. a.

** Obtained using the profile criterion described in Section IV. B. 2. b.

Table 9. Comparisons for Example INDIVIDUAL RAYS (See Table 8 for symbols)

Ray height (km) (Selected from original tabulated profile)	Cumulative		ϵ (mrad)		l (m)		L (km)	
	Layer A	Count C	A	A - C	A	A - C	A	A - C
.000	0	0						
.952 (1st)	30	3	5.112	.044	2.352	-.008	127.875	.280
2.007	40	7	7.304	.045	7.152	-.010	185.259	.289
5.381	56	17	10.221	.047	22.404	-.015	300.796	.294
20.105 (last)	83	59	12.998	.047	70.142	-.016	564.402	.294
300.	103	80	13.208	.048	80.626	-.170	2054.909	.300
Data above for warm sector profile, $\hat{h} = 0$ km								
2.0	0	0						
2.163 (1st)	16	2	2.123	-.030	.169	-.0007	52.893	-.134
2.793	31	4	3.867	-.028	1.369	-.003	114.872	-.146
6.816	56	16	7.046	-.025	14.194	-.007	275.400	-.153
23.377 (last)	90	66	9.507	-.025	55.422	-.008	565.052	-.159
300.	109	87	9.628	-.025	61.879	-.097	2027.367	-.159
Data above for post cold front profile, $\hat{h} = 2$ km								

The differences in ℓ are due primarily to the differences in the two different interpolated profiles between the last tabulated height and the assumed satellite height. The more widely spaced layers of method C result in systematically greater N values at all heights and yield slightly larger ℓ values. These ℓ differences are probably inconsequential for applications but well within the resolution of the tracing methods.

In conclusion, we note that relatively minor differences in the specification of a profile for ray-tracing result in significant differences in the refraction variables, especially the path length L. This is important when considering the differential refraction of rays travelling at different heights in the same profile or when studying variations in refraction of similar rays in different profiles. Method C is superior in that it specifies exactly the same profile for any ray. Method A gives "exact" results for a profile which is exponential between tabulated points by using as many layers as are required to obtain precise refraction values.

d) Internal Consistency Checks. The programs described above determine ϵ , ℓ and L using essentially separate calculations according to equations (2.18), (2.19) and (2.20). These quantities are related by equations (1.1) through (1.6), so that a number of consistency checks are possible. The quantities, r , θ , α , ϵ , ℓ and L are calculated at each step in the trace. Denoting the final values at satellite height with subscript s, we expect:

$$\hat{n} \hat{r} \tan \alpha_s = r_s \sin \alpha_s ; \quad (4.5a)$$

$$\hat{n} \hat{r} (\epsilon_s + \tan \alpha_s) \doteq r_s \left[\left(1 + \frac{1}{2} \epsilon_s^2\right) \sin \theta_s - \frac{1}{3} \epsilon_s \cos \theta_s \right] \equiv T_s ; \quad (4.5b)$$

$$T_s + \ell_s \doteq L_s . \quad (4.5c)$$

This procedure also provides a check on roundoff errors. It was found that using RAYTRACE A and the 10^{-4} criterion, L_s (equation (4.5c)) often failed to agree by many meters. With 10^{-5} criterion or RAYTRACE C, the terms agreed within a few meters for calculations using single precision on an IBM 360/65. The lack of complete agreement is attributed to roundoff errors. Subsequent calculations using the same program on a Univac 1108 produced in all cases agreement within one meter.

3. Ray-tracing With Specified Minimum Ray Heights

The minimum ray height \hat{h} is sufficient to uniquely define a ray in a given spherically stratified medium. The basic results obtained by tracing a single ray are the coordinates (r, θ) and the refraction variables ϵ , ℓ and L at points along the ray. The "total" values ϵ_s and ℓ_s at satellite height are independent of (r_s, θ_s) due to the vanishing N-gradient at great heights. Figures 11a - d show examples of the distribution of refraction effects along a few selected rays. These results generally confirm the analytical predictions derived earlier (Sargeant 1968). However, the "weighting functions" of Figs.

11c, d are no longer smooth but reflect the occurrence of sharp-gradient layers which are absent in purely exponential profiles.

The quantities $\epsilon_S(\hat{h})$ and $l_S(\hat{h})$, once derived for a given profile, suffice to determine the total optical path $L_S(\hat{h})$ for an arbitrary satellite height r_S (or angular position θ_S) without any further ray-tracing. Equation (1.1) determines one of the pair (r_S, θ_S) in terms of the other when $\epsilon_S = \theta_S - \alpha_S$ is known. Then L_S may be calculated using equations (1.4) and (1.5) and l_S . Table 10 presents results obtained for the various Campbell Island profiles using ray-tracing method C described above. All profiles were adjusted to their typical surface pressure as described in Section III. B. 1, and an "average" profile derived from values interpolated to half-kilometer intervals was also traced.

Example values of θ_S , L_S and $L_{eS} = L_S - r_S \sin \theta_S$ are also included for the special case $h_S = 300$ km. The quantity L_{eS} can be interpreted as half the optical path length in excess of the direct, geometrical distance between two satellites at 300 km.

The range of variation of refraction variables for these profiles is illustrated in Figs. 12a, b, and c. As the results are so similar in absolute value at corresponding heights, some examples of the deviations from the average results are shown in Figure 13 along with the deviations of the corresponding profiles from the average profile. (The "average results" do not include those for the average profile. It is found, incidentally, that the deviations of results (ϵ , etc.) obtained for individual profiles from those obtained for the average profile behave rather differently from Figs. 13a, b, and c, but this is not shown here.) It is apparent from these figures that N-profile deviations and refraction deviations are closely related, as expected. Also, because gradients are most important, the relation is not a simple, point-for-point linear correlation. This question deserves additional attention, and is discussed further in Section V.

In any case, such deviations (at a given \hat{h}) are certainly well within the range of expected measurement resolution. Unfortunately, the same pair of satellites will not observe rays with the same minimum ray heights in these different profiles, as can be observed from the differing θ_S values for the same \hat{h} in Table 10. In general, a ray which is bent more will go further before reaching the satellite height. Thus, to traverse the same angular distance θ_S , observed rays will tend to pass at heights which yield comparable bending angles ϵ_S . The extent of this "compensation" will be investigated in the next section.

4. Ray-tracing with Specified Satellite Positions

It is not possible to specify (r_S, θ_S) and \hat{h} independently for a given N-profile. In fact, \hat{h} alone determines ϵ_S and l_S while \hat{h} and r_S or θ_S determine L_S . More simply, a ray launched at a given point with a given direction will go wherever refraction takes it. In a spherically strati-

Table 10. Ray-tracing Results for Selected Minimum Ray Heights

^A h (km)	RIDGE PROFILE						WEST OF RIDGE PROFILE				
	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)	
0.0	13.180	80.020	253.62	2054.732	313.293	13.600	79.744	264.66	2057.379	313.709	
0.5	12.981	75.014	243.28	2052.224	312.900	13.057	74.796	245.04	2052.719	312.978	
1.0	16.626	69.810	346.72	2074.187	316.348	15.263	69.850	302.91	2065.588	314.998	
1.5	9.197	66.413	150.45	2026.040	308.789	9.995	66.233	165.57	2031.059	309.577	
2.0	8.272	62.857	130.72	2018.830	307.657	8.424	62.671	133.07	2019.766	307.804	
3.0	7.391	56.229	110.28	2010.452	306.342	7.457	55.912	110.94	2010.877	306.408	
4.0	6.726	50.107	94.78	2003.399	305.234	6.718	49.780	94.35	2003.360	305.229	
6.0	5.486	39.340	68.93	1989.754	303.093	5.492	39.019	68.68	1989.811	303.102	
8.0	4.460	30.399	49.88	1977.327	301.144	4.434	30.107	49.36	1977.182	301.121	
12.0	2.855	16.610	24.53	1955.048	297.650	2.783	16.524	24.05	1954.596	297.579	
16.0	1.398	8.991	10.87	1933.214	294.228	1.419	8.873	10.81	1933.349	294.249	
20.0	0.755	4.928	5.47	1916.139	291.553	0.750	4.805	5.34	1916.109	291.549	
24.0	0.400	2.716	2.87	1900.646	289.128	0.392	2.618	2.76	1900.596	289.120	

327

Table 10 continued on next page

Table 10 (continued)

\hat{h} (km)	WARM SECTOR PROFILE					POST COLD FRONT PROFILE				
	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)
0.0	13.208	80.626	254.98	2054.904	313.321	12.139	78.852	225.96	2048.354	312.292
0.5	12.350	75.874	228.08	2048.239	312.274	11.441	74.298	204.81	2042.669	311.400
1.0	11.590	71.340	205.19	2042.168	311.321	10.875	69.925	187.69	2037.792	310.634
1.5	11.082	67.019	189.26	2037.709	310.621	10.542	65.725	176.27	2034.406	310.102
2.0	9.908	63.068	160.58	2029.010	309.255	9.628	61.879	153.94	2027.362	308.996
3.0	7.665	56.079	114.23	2012.153	306.609	7.523	55.105	111.86	2011.628	306.526
4.0	6.993	49.788	98.08	2005.070	305.497	6.773	48.945	94.24	2003.734	305.287
6.0	5.506	38.863	68.67	1989.930	303.121	5.540	38.145	68.33	1990.149	303.156
8.0	4.393	30.044	48.94	1976.950	301.084	4.510	29.117	49.04	1977.722	301.206
12.0	2.772	16.598	24.06	1954.541	297.570	2.530	16.062	22.28	1953.066	297.339
16.0	1.382	9.083	10.92	1933.125	294.214	1.375	8.889	10.71	1933.078	294.207
20.0	0.717	5.114	5.61	1915.904	291.517	0.740	4.903	5.43	1916.051	291.540
24.0	0.398	2.989	3.14	1900.626	289.125	0.379	2.761	2.90	1900.514	289.107

Table 10 continued on next page

Table 10 (continued)

λ h (km)	EAST OF RIDGE PROFILE					AVERAGE PROFILE				
	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)
0.0	12.641	79.416	239.02	2051.454	312.779	12.934	79.742	246.90	2053.254	313.061
0.5	12.084	74.692	220.39	2046.674	312.029	12.385	74.947	228.02	2048.529	312.320
1.0	12.230	70.017	219.16	2046.340	311.976	13.002	70.237	238.96	2051.225	312.743
1.5	9.850	66.148	162.61	2030.148	309.434	9.974	66.334	165.24	2030.866	309.546
2.0	8.585	62.547	135.67	2020.804	307.967	8.855	62.618	140.43	2022.472	308.229
3.0	7.499	55.846	111.49	2011.142	306.450	7.504	55.826	111.56	2011.178	306.456
4.0	6.727	49.710	94.39	2003.431	305.240	6.786	49.655	95.13	2003.798	305.297
6.0	5.441	39.026	68.13	1989.496	303.053	5.470	38.871	68.29	1989.689	303.083
8.0	4.435	30.133	49.39	1977.184	301.121	4.441	29.949	49.26	1977.242	301.130
12.0	2.796	16.381	23.97	1954.700	297.595	2.738	16.427	23.71	1954.338	297.538
16.0	1.381	8.901	10.74	1933.119	294.213	1.388	8.934	10.79	1933.158	294.219
20.0	0.735	4.910	5.43	1916.018	291.534	0.743	4.911	5.44	1916.067	291.542
24.0	0.382	2.777	2.92	1900.528	289.110	0.389	2.740	2.88	1900.581	289.118

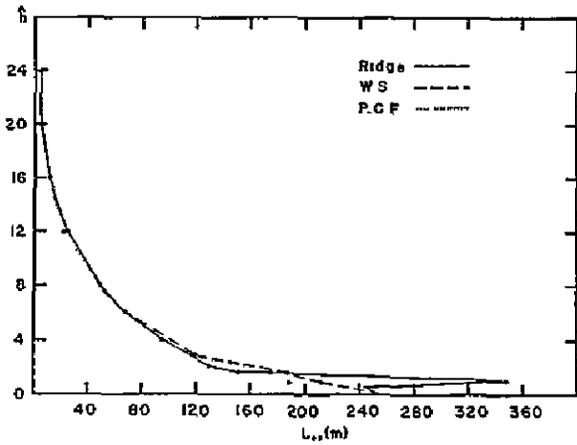
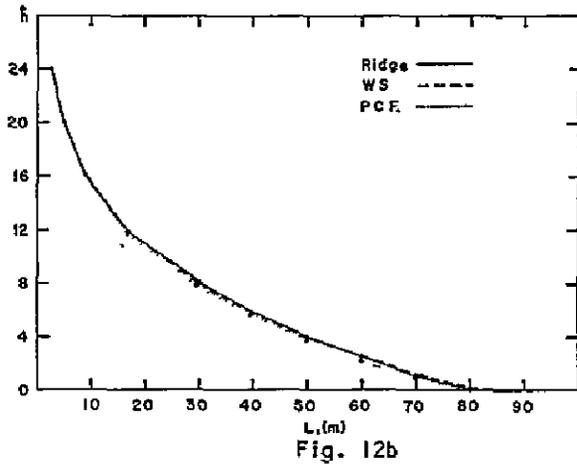
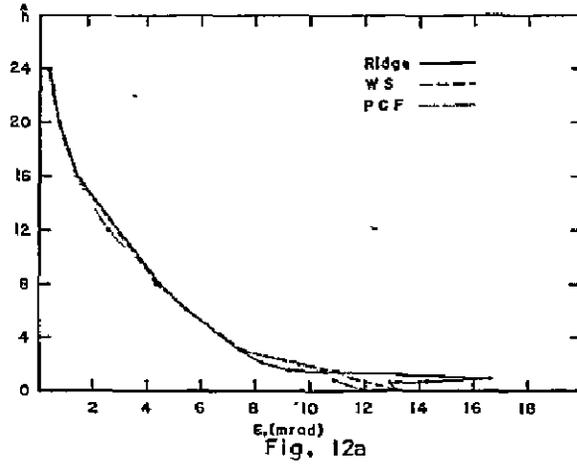


Fig. 12a - c. Range of variation of refraction variables, ϵ_s , L_s , L_{es} for selected "real" refractivity profiles. See Fig. 2.

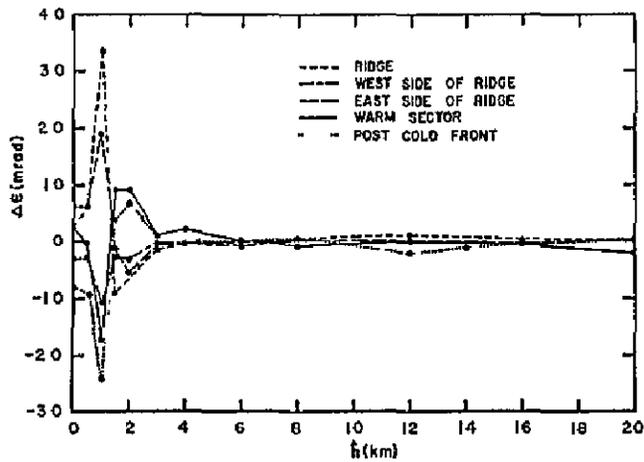


Fig. 13a

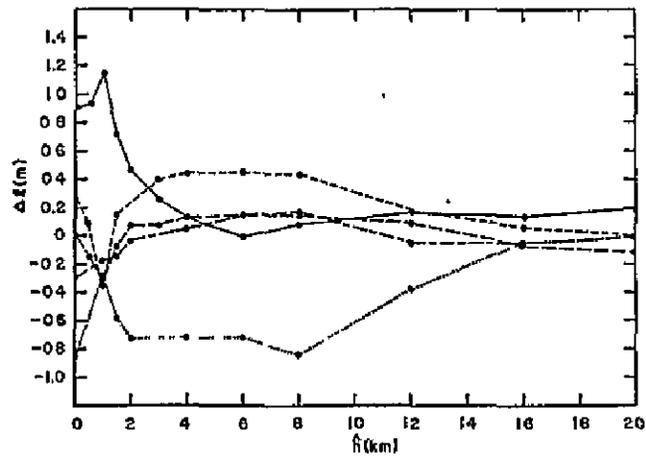


Fig. 13b

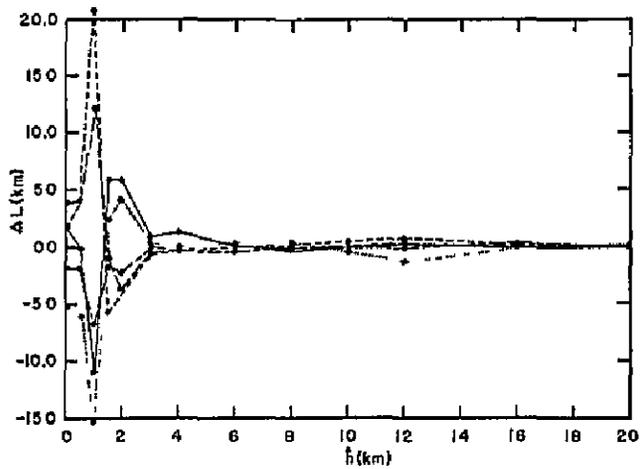


Fig. 13c

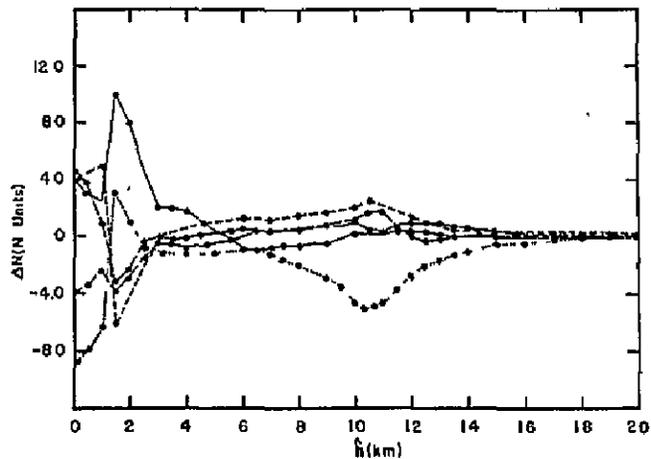


Fig. 13d

Fig. 13a - c. Deviations of refraction variables from "average" results.

Fig. 13d Deviations of the various profiles from the "average" profile.

fied medium with moderate gradients (weaker than "ducting"), the ray will eventually reach any great height, but only after traversing a uniquely specified central angle.

Furthermore, given (r_s, θ_s) and a profile, \hat{h} is not necessarily even uniquely determined. That is, there can be "multipath," i. e., multiple rays through the same points, if the refractivity profile is appropriate. The Campbell Island profiles provide some examples as can be seen from the results already presented. Table 10 yields samples of the function $\hat{h}(\theta_s)$ for a fixed $\hat{h}_s = 300$ km. Figure 14 illustrates an example. Note that multiple rays are possible for Ridge and probably other profiles when θ_s is between 313 and 316 mrad. This results, of course, from the steep N-gradients at heights near 1 km.

In order to reasonably simulate the refraction observations anticipated for occultation satellites, it is necessary to prescribe the satellite configuration, viz. h_s and θ_s , a priori. Then \hat{h} and the refraction variables for the ray or rays which intersect (r_s, θ_s) must be determined for given profiles. A ray-tracing method requires some iterative procedure equivalent to launching a "trial" ray toward the satellite, re-aiming a new ray based on the result, etc. The problem is to find an efficient, rapidly converging algorithm which minimizes the required number of ray-traces.

The following technique has been devised. First, we recall that for any ray (assuming $n(r_s) = 1$):

$$r_s \cos \alpha(r_s) = \hat{n} \hat{r}; \quad (4.6)$$

$$\theta(r_s) = \alpha(r_s) + \epsilon_s. \quad (4.7)$$

For the desired ray, $\theta(r_s) = \theta_s$, where θ_s is the prespecified angular separation. We, therefore, form the function

$$\phi(r_s) \equiv \epsilon_s + \alpha(r_s) - \theta_s \quad (4.8)$$

and seek to relax ϕ to zero.

Let the desired value of \hat{h} be denoted \hat{h}_s . Now, given any initial guess for \hat{h}_s , say \hat{h}_1 , we can calculate $\epsilon_s(\hat{h}_1)$ by ray-tracing, $\alpha_1(r_s)$ by equation (4.6), and $\phi_1(r_s)$ by equation (4.8). We have thereby implicitly defined the function (not necessarily single-valued) $\hat{h}(\phi)$, and we seek $\hat{h}(0)$. Given two values $\hat{h}(\phi_1)$ and $\hat{h}(\phi_2)$, we can approximate $\hat{h}(\phi)$ by the straight line $p_2(\phi)$ and estimate $h_3 = p_2(0)$. More generally, given k values for $\hat{h}(\phi)$, we can form a polynomial $p_k(\phi)$ which assumes those values and estimate $\hat{h}_{k+1} = p_k(0)$. This procedure is attractive because it utilizes the information from all of the previous cumbersome ray-traces.

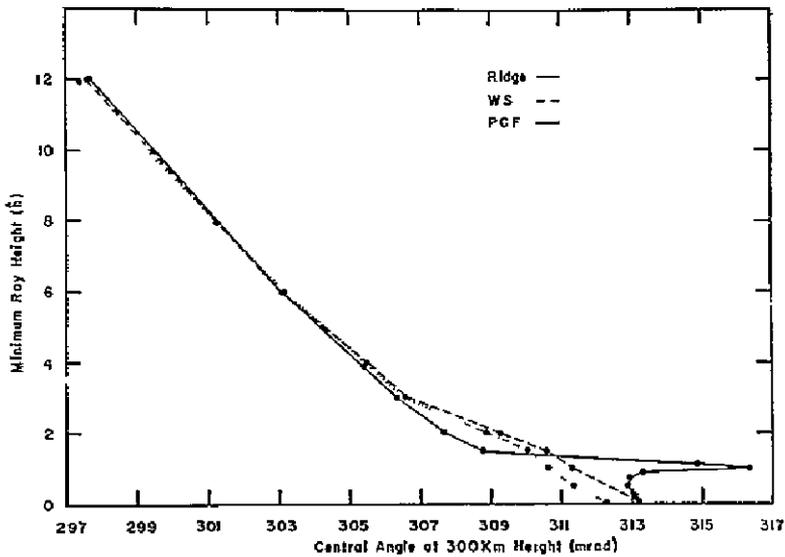


Fig. 14. Comparison of profiles with (Ridge) and without (W. S. and P. C. F.) multiple ray paths.

The evaluation of a polynomial $p_k(\phi)$ which takes on k specified values is the classical Lagrangian interpolation problem. One form (Hildebrand, 1956) is

$$p_k(\phi) = \left[\sum_{j=1}^k \frac{h(\phi_j)}{(\phi_j - \phi)\Phi_j} \right] / \left[\sum_{j=1}^k \frac{1}{(\phi_j - \phi)\Phi_j} \right]$$

where $\Phi_j = (\phi_1 - \phi_j)(\phi_2 - \phi_j) \dots (\phi_{j-1} - \phi_j)(\phi_{j+1} - \phi_j) \dots (\phi_k - \phi_j)$. Thus,

$$p_k(0) = \sum_{j=1}^k \frac{h_j}{\phi_j \Phi_j} / \sum_{j=1}^k \frac{1}{\phi_j \Phi_j} \quad (4.9)$$

The Φ_j 's can be computed easily from earlier ones (smaller k) by a simple iterative procedure.

Although the above procedure is elegant and effective when $\hat{h}(\phi)$ is well-behaved (i. e., when the profile is reasonably smooth), troublesome cases have been encountered. In case of multipath, it is desirable to "localize" the \hat{h} estimates because of the multiple zeros of $\phi(\hat{h})$. One procedure which proved successful was to linearly interpolate (or extrapolate) the current two "best" (smallest ϕ) estimates of \hat{h} , retaining one estimate for each sign of ϕ after these were obtained.

Either of the above procedures requires a convergence criterion. The method selected required successive values of \hat{h} to agree within one meter. This should normally define ϵ_S and θ_S within one microradian. (However, even one microradian in θ_S corresponds to about six meters in L_S , so that "small" errors can be serious.) Since the profiles themselves are specified in layers much thicker than one meter, it was considered extravagant to iterate ray-traces with \hat{h} values so nearly equal, even though only ϵ_S values were being calculated without detailed print-out. Consequently, whenever two successive traces yielded ϵ_S values within 0.1% of one another, tracing was terminated and additional ϵ_S values were obtained by interpolation. Then after convergence to \hat{h} , a complete trace was performed. As it turned out, these profiles were sufficiently "nasty" so that this procedure produced convergence using interpolated ϵ_S values which were noticeably different from the finally traced values in several cases. In addition, the procedure converged to a spurious \hat{h} in one case, and not at all in two other cases because of a logical error in the program. (The program, left in Australia, was not rerun.)

Some illustrative results are presented in Table 11. Note that the large deviations in ϵ observed between profiles for the same \hat{h} (Table 10) are largely compensated as anticipated above. Even more striking is the near equality of the excess path length values L_{eS} , which agree within about one meter whenever the θ_S values are reasonably close. Note also the great advantage in being able to calculate L_{eS} accurately from ϵ_S , l_S and slightly inaccurate values for θ_S . The differences in the traced values of L_S for nearly equal θ_S are almost entirely explained by the expected differences in $L_S - L_{eS} = r_S \sin \theta_S$ due to θ_S differences.

V. THE INVERSION PROBLEM

A. Introduction

Suppose that accurate propagation measurements can somehow be made. How can desirable atmospheric information be extracted from the data? We have seen that microwave refraction variables are quite sensitive to the vertical profile of refractivity, and especially to strong gradients. For dry air, refractivity is merely proportional to density, so that the refractivity profile determines the mass distribution. However, the contribution of water vapor to refractivity is disproportionate because of the polar molecules. This fact complicates the inference of density, but potentially provides information on the humidity profile. Can we extract this information?

As the physical relations governing the propagation depend on atmospheric structure through refractivity, one might suppose that a reasonable inversion procedure should seek to infer refractivity from refraction measurements. The

Table 11. Ray-tracing Results for Selected Satellite Positions

Profile	\hat{h} (km)	ϵ_s (mrad)	l_s (m)	L_{es}^* (m)	L_s (km)	θ_s (mrad)
Nominal $\theta_s = 300$ mrad						
Ridge	9.370	3.961	25.123	40.44	1970.030	299.999
West of Ridge	9.347	3.944	24.932	40.13	1970.014	299.997
Warm Sector	9.266	3.905	25.238	40.13	1970.030	299.999
Post Cold Front	9.457	3.985	23.524	39.03	1970.029	299.999
East of Ridge	9.331	3.939	25.029	40.18	1970.029	299.999
Nominal $\theta_s = 304$ mrad						
Ridge	5.170	6.015	43.566	79.20	1995.526	303.999
West of Ridge	5.145	6.001	43.374	78.85	1995.526	303.999
Warm Sector	5.298	6.070	42.433	78.72	1995.529	304.000
Post Cold Front	5.183	6.011	42.330	77.92	1995.511	303.997
East of Ridge	5.118	5.987	43.488	78.80	1995.529	303.999
Nominal $\theta_s = 306$ mrad						
Ridge	3.318	7.188	54.225	105.32	2008.272	305.999
West of Ridge	3.355	7.203	53.670	104.98	2008.271	305.999
Warm Sector	3.515	7.281	52.776	105.18	2008.270	305.999
Post Cold Front	3.374	7.208	52.739	104.11	2008.268	305.999
East of Ridge	3.387	7.217	53.398	104.90	2008.271	305.999

* L_{es} calculated using equation (1.5).

Table 11 (continued)

Profile	\hat{h} (k _m)	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)
Nominal $\theta_s = 307$ mrad						
Ridge	2.413	7.789	60.062	120.17	2014.637	306.998
West of Ridge	2.676	7.909	58.012	119.96	2014.643	306.999
Warm Sector	2.834	7.984	57.167	120.30	2014.644	306.999
Post Cold Front	2.727	7.928	56.882	119.13	2014.643	306.999
East of Ridge	2.498*	7.905*	59.144*	121.06*	2015.128*	307.075*
Nominal $\theta_s = 308$ mrad						
Ridge	1.187	8.558	63.752	136.43	2021.003	307.997
West of Ridge	1.826*	8.644*	63.912*	138.06*	2021.640*	308.097*
Warm Sector	2.523	8.859	59.250	137.08	2021.012	307.999
Post Cold Front	2.182	8.693	60.616	135.58	2020.965	307.992
East of Ridge	1.984	8.610	62.660	136.22	2021.007	307.998
Nominal $\theta_s = 309$ mrad						
Ridge	1.383*	11.678*	67.150*	203.02*	2042.082*	311.306*
West of Ridge	1.567*	9.672*	65.723*	158.73*	2028.847*	309.229*
Warm Sector	2.110	9.695	62.233	155.58	2027.384	308.999
Post Cold Front	1.999	9.630	61.884	153.99	2027.380	308.998
East of Ridge	1.577	9.445	65.598	154.24	2027.387	309.000

* NOTE: Not the desired θ_s value.

Table 11 (continued)

Profile	\hat{h} (km)	ϵ_s (mrad)	l_s (m)	L_{es} (m)	L_s (km)	θ_s (mrad)
Nominal $\theta_s = 310$ mrad						
Ridge	1.438	10.385	66.836	174.13	2033.756	309.999
West of Ridge	1.398*	10.624*	66.982*	179.31*	2035.306*	310.243*
Warm Sector	1.727	10.547	65.187	175.80	2033.749	309.999
Post Cold Front	1.554	10.459	65.293	174.10	2033.750	309.999
East of Ridge	1.390*	10.583*	66.958*	178.41*	2035.081*	310.208*
Nominal $\theta_s = 311$ mrad						
Ridge	1.400	11.376	67.054	195.93	2040.126	311.000
West of Ridge	1.326	11.352	67.490	195.83	2040.090	310.994
Warm Sector	1.253	11.368	69.108	197.80	2040.128	311.000
Post Cold Front	0.767	11.146	71.937	195.73	2040.118	310.999
East of Ridge	-	-	-	-	-	-
Program Error						
Nominal $\theta_s = 312$ mrad						
Ridge	1.344	12.360	67.396	219.70	2046.488	311.998
West of Ridge	1.169*	12.711*	68.663*	229.84*	2049.066*	312.403*
Warm Sector	0.652	12.134	74.467	221.32	2046.491	311.999
Post Cold Front	0.183	11.916	77.158	218.86	2046.488	311.999
East of Ridge	-	-	-	-	-	-
Program Error						

*NOTE: Not the desired θ_s value.

problem of inferring density, pressure, temperature, humidity, etc. could then be separated entirely, in view of equation (2.26). On the other hand, if density, temperature, etc. are desired, perhaps a reasonable procedure should relate these directly to the observed variables.

Perhaps a more appropriate question is what can be inferred from refraction measurements? There are at least two ways to approach this question: they might be called "mathematical" and "statistical." A theoretically direct and satisfying approach is to begin with a physical relation such as equation (1.3). This can be considered to be an integral equation relating the functions $\epsilon(r)$ and $n(r)$. An analogous equation relates $L(r)$ and $n(r)$. The "mathematical" problem is then the solution of integral equations such as these. For example, we should like to know the necessary and sufficient conditions for a solution. We must inquire whether or not there are analytical or numerical techniques available to deduce $n(r)$. If so, how sensitive are these methods to errors in the measurements or in the theoretical formulation itself? How many measurements of which variables are required to determine useful information regarding $n(r)$? Is there any optimum choice?

Physical insight aided by recent experience with an analogous problem, that of inferring vertical temperature profiles from infrared radiance spectra, suggests caution when applying the mathematical approach. Practical measurements may very well not provide a unique solution to the problem posed above or may yield a highly unstable solution, especially if an inherently nonlinear problem is linearized. Direct mathematical inversion of the integral equations assumes that we know very little about the nature of the desired solution, aside from the existence of necessary derivatives, etc. In fact, we know a great deal about the desired solution; most of the "possible" profiles (i. e., consistent with propagation laws) which correspond to a given set of measurements may be discarded a priori. We therefore seek an efficient procedure which incorporates our prior knowledge to better specify the problem.

One can also view the inversion problem as one of statistical estimation of parameters. Given a set of measured "refraction variables," what are the probable values of a set of related "atmospheric variables." Here we explicitly recognize that the two sets of variables are not only mathematically related (e. g., equations (1.3) and (2.26)) but also statistically related, both within each set and between sets. Both kinds of relations have physical content. From the statistical point of view, we need to establish which variables should be included in each set, how "prior statistical knowledge" should be introduced, and what sort of criteria should be applied to specify the "probable" values of "inferred" variables when the "observed" variables are given.

Both the "mathematical" and the "statistical" approaches have been applied to some extent, both for the occultation problem and for the analogous infrared problem. The author feels that the prospects for useful occultation inversions

are hopeful, but that the problem requires a good deal more study. A brief review of some of the efforts to date is presented below.

B. General Considerations

Analysis of the inversion problem depends on the specification of a) the meteorological variables to be determined—including accuracy and sampling requirements, b) the propagation variables to be measured—including expected resolution and accuracy, c) the configuration of the satellites—including accuracy of location, and d) the nature and accuracy of complementary data to be available—including both concurrent and historical data. Of course, these factors cannot be specified independently, and the specifications are continually changing because of scientific, technological and operational developments. This problem was discussed briefly in Section I. B.

The principal meteorological variables we can hope to infer are the vertical distribution of density and water vapor and with these also the vertical profiles of temperature and pressure. More specialized applications, such as determination of pressure-height reference levels are also likely, as mentioned in Section I. B. The principal candidates for microwave measurements are the phase path length (relative and absolute), refraction angle or angle of arrival, signal amplitude, and dispersion between selected frequencies; all as a continuous function of time or position, and all measured over a path between one or more pairs of satellites. The principal satellite configuration proposed consists of a mother and several daughters in the same circular orbit. In this case the actual separation between pairs remains essentially constant, but only a few limited height intervals are effectively sampled. Other configurations and variables are conceivable (and have been proposed) but are not considered here.

The preliminary report (Sargeant, 1968) includes calculations of the sensitivity of the refraction variables ϵ , l and L to variations in refractivity, scale height, and minimum ray height for propagation in an exponential atmosphere. The sensitivity and error requirements do not appear technically unreasonable, except that ϵ may not be measurable accurately enough to add much information to the phase path L alone. The minimum ray height \hat{h} also must be determined from the measurements. The satellite altitudes must be known as accurately as the required resolution in \hat{h} , and the satellite separations must be known as accurately as the required resolution in L .

Requirements for useful attenuation measurements and bandwidth to accommodate expected scintillations are discussed by Pomalaza (1969). Considerations for dispersion measurements have been investigated by Graves and Fischbach (1969) and others.

All of the proposed microwave variables describe effects integrated along the path, and thus constitute weighted spatial averages. A "universal" set of

ideal weighting functions was derived in the preliminary report, and specific realistic examples are given in Section I. V. B. 3. Comparison with similar functions derived for passive infrared radiance sounding reveals some significant differences. In occultation, the vertical region of contribution is sharply bounded below by the minimum ray height, and is generally thinner than for the infrared case. The degree of overlap between the various measurements is thereby considerably less. Also, the microwave measurements represent a much longer horizontal average (along the path). Despite these differences, the inversion problems are basically similar. What can be determined is a probable profile of refractivity (or other variable) and not independent values at particular heights.

C. Mathematical Inversion

The application of analytically-based integral-inversion techniques has been vigorously propounded by Graves and Fischbach (1969). They point out that the refraction angle integral (equation (1.3)) has an Abelian inverse giving the refractivity profile. However, there are difficulties in application to the practical microwave problem. First, the refraction angle must be known as a function of the "impact parameter" ($\hat{n}\hat{r}$)—which is not the case. Second, the entire continuous function $\epsilon(\hat{n}\hat{r})$ is required, whereas only a few samples are obtained using the proposed configuration. Interpolation techniques provide a possible solution to the latter problem.

In any case, as the authors point out, the analytical technique provides a potentially powerful tool for feasibility studies and particularly for error analyses.

The cited report includes an extensive bibliography relevant to the "mathematical" inversion problem.

D. Statistical Inversion

This method is basically parameter determination. In its simpler forms, one might call it "model fitting." Such a method has been adopted by Lusignan, et al. (1969), and Pomalaza (1969). As applied by the Stanford group, the excess phase path length (herein L_{es}) is assumed to be an approximately linear function of suitable profile parameters, and least-mean-square values for the parameters are determined by an iterative procedure. Of course, the sufficiency of such a method depends on the identification of suitable parameters. Empirical orthogonal functions have been used to represent both temperature and humidity profiles in successful inversions of simulated microwave data derived by ray-tracing using radiosonde profiles.

Pomalaza concludes from his study that useful temperature and pressure profiles can be obtained above about 360 mb from phase measurements alone

(neglecting water vapor entirely), down to about 500 or 600 mb by using water vapor climatology, and down to 700 or 800 mb by adding attenuation measurements in the 1.35 cm water vapor absorption line. Accuracies of about 0.3 m in phase path length and about 0.5 dB in attenuation are required using a configuration of 5 or 6 satellites in order to define temperature and pressure profiles with rms errors of about 2°K and 3 mb, respectively.

Although these studies demonstrate that practical refraction measurements probably can be inverted to provide useful meteorological profiles, the techniques used are by no means optimal. There is considerable flexibility available in choosing the number and separations of the satellites, for example, as well as the microwave measurements to be made. Probably more important, a great variety of meteorological parameterizations are conceivable, and various combinations of historical and concurrent meteorological data could be employed. Some incisive studies into these alternatives might well be expected to yield substantial improvements in the derived meteorological variables, particularly as regards water vapor in the lower troposphere.

Smith (1970) has shown that linear regression techniques can be used to perform "inversions" of infrared vertical-sounder data to obtain useful temperature profiles throughout the troposphere, even though clouds distort or completely obscure the lower region. The inherent correlation between meteorological variables at various heights should also assist the inversion of microwave data. As noted earlier, the combination of both infrared data and occultation data may result in useful profiles of both temperature and humidity. Some simulation studies along this line would seem appropriate.

One interesting problem is the selection of a more optimal collection of microwave and meteorological variables, and the development of techniques for extracting more nearly maximal meteorological information from the measurements. Multivariate statistical analyses may prove applicable to this problem as well. A study is already underway to investigate the utility of canonical correlation analysis for this problem.

One cannot escape the feeling, however, that linear statistical analyses constitute a "dumb-John" approach to the inversion problem. Just as integral-inversion techniques ignore much of what we know about expected solutions because of meteorological constraints (often expressed only by climatology rather than equations), linear correlation analyses ignore much of what we know about the physical relations between and among propagation variables and meteorological variables. In particular, we know that many of the relations are nonlinear. The analogous situation for infrared profile inversion was pointed out by Chahine (1969), when he developed a stable, iterative nonlinear inversion procedure for a class of radiative transfer problems. This method has recently been extended by W. L. Smith (private communication, 1970) and applied to the SIRS data reduction. Perhaps a truly optimal solution must make

use of both nonlinear, physically-based relations and convenient statistical relations.

In conclusion, it is noted that current views of observational requirements for numerical modeling explicitly recognize that ordinary meteorological variable are highly redundant because of dynamical and other physical constraints. Present primitive-equation models constitute an elaborate four-dimensional scheme to express our best, tractable statement of these relations, and these relations are numerically displayed in all four dimensions during the course of the calculations. Perhaps it is useful to apply such a view to potential observational techniques and variables. That is, perhaps all of the available four-dimensional data, the meteorological relations (models), and tractable physical expressions for the probing process should be applied to extract the maximum "new" information from a "measurement." Thus, in a prediction process, the difference between an actual measurement and the result expected from deterministic "prediction" of the atmosphere and the observational event can be considered as a fine vernier with which to adjust predicted atmospheric variables.

The implementation of such a vision stands as a challenge for meteorologist and technologist alike.

APPENDIX A

ANALYTICAL CALCULATION OF LAGUERRE COEFFICIENTS

From equation (3.4), $b_k = \frac{1}{(k!)^2} \int_0^\infty N(Hx) L_k(x) dx$.

Theorem: $\int_0^\infty e^{-x/\beta} L_k(x) dx = k! \beta (1 - \beta)^k$.

This can be proved by induction on k using the recursion relations (equations (2.7)).

A. Exponential Profile: $N(h) = N_0 e^{-h/H^*}$

$N(Hx) = N_0 e^{-x/\beta}$ where $\beta = H^*/H$

$b_k = (N_0/k!) \beta (1 - \beta)^k$.

Note: Define $\tilde{N}_k(h) = e^{-h/H} \sum_{R=0}^{m-1} b_k L_k(h/H)$,

$$\text{then } \mathcal{N}(0) = \sum_{k=0}^{m-1} k! b_k = N_0 [1 - (1 - \beta)^m].$$

B. Bi-exponential Profile: $N(h) = N_1 e^{-h/H_1} + N_2 e^{-h/H_2}$;

$$N(Hx) = N_1 e^{-x/\beta_1} + N_2 e^{-x/\beta_2} \text{ where } \beta_1 = H_1/H, \beta_2 = H_2/H;$$

$$b_k = \frac{1}{Hk!} \left[N_1 H_1 \left(1 - \frac{H_1}{H}\right)^k + N_2 H_2 \left(1 - \frac{H_2}{H}\right)^k \right].$$

$$\text{Also, } \mathcal{N}(0) = N_1 \left[1 - \left(1 - \frac{H_1}{H}\right)^m\right] + N_2 \left[1 - \left(1 - \frac{H_2}{H}\right)^m\right].$$

Note that: $\mathcal{N}(0) = N(0)$ if m is chosen odd and

$$H = \frac{H_1 + aH_2}{1 + a}, \text{ where } a = \left(\frac{N_2}{N_1}\right)^{1/m}.$$

A tedious calculation also reveals that the minimum mean-square error term ξ_{\min}/H (see equation following (3.4)) is approximately minimized (for large m) by the above choice of H . This particular choice is called "optimum" in the text.

APPENDIX B

CALCULATION OF THE CONTRIBUTION TO THE LAGUERRE COEFFICIENTS

DUE TO THE UPPER PORTION OF THE RAY PATH

The contribution above a height h_u is given by (see equation (3.4))

$$\Delta b_k = \frac{1}{(k!)^2} \int_u^\infty N(Hx) L_k(x) dx \text{ where } u = h_u/H.$$

If $N(h) = N_u e^{-\left(\frac{h-h_u}{H}\right)}$ above height h_u , then

$$\Delta b_k = \frac{N_u}{(k!)^2} \int_0^\infty e^{-t} L_k(u+t) dt \equiv \frac{N_u}{(k!)^2} I_k.$$

Theorem: $I_{k+1} = kI_k - uL_k(u)$; $I_0 = I$.

This can be proved by induction on k using recursion relations (equation (2.7)) and integration by parts.

The "corrections" Δb_k can thus be calculated by a simple recursive computer program.

ACKNOWLEDGMENTS

The writer gratefully acknowledges V. E. Suomi for stimulating this work, the international Antarctic Meteorological Research Centre and Staff for graciously hosting and assisting me during most of the work, The Australian Commonwealth Bureau of Meteorology for providing their computing facilities, Mary Adams and Margaret Bell for extensive programming assistance, and D. W. Thomson and J. A. King for assistance in preparing the report.

The research was supported by the National Science Foundation under Grant GA-1167, and the preparation and publication of the report by the National Aeronautics and Space Administration under Grant NAS 5-11542.

REFERENCES

- Abramowitz, Milton, and Irene A. Stegun, ed., Handbook of Mathematical Functions, AMS 55, U. S. Government Printing Office, 1964.
- Bean, B. R., and E. J. Dutton, Radio Meteorology, N.B.S. Monograph 92, U. S. Government Printing Office, 1966.
- Chahine, Moustafa T., Inverse Problems in Radiative Transfer: Determination of Atmospheric Parameters, submitted for publication to J. Opt. Soc. Amer., 1969.
- Graves, M. E. and F. F. Fischbach, "Analysis of Microwave Occultation Techniques for Atmospheric Soundings," Univ. of Michigan ORA Tech. Rept. 05863-16-T, Jan. 1969.
- Hildebrand, F. B., Introduction to Numerical Analysis, McGraw-Hill Book Co., New York, 1956.
- Lusignan, B., private correspondence, 1970.

- Lusignan, Bruce, Gary Modrell, Angus Morrison, Jose Pomalaza and Steven Ungar, "Sensing the Earth's Atmosphere with Occultation Satellites," Proc. I. E. E., Vol. 57, 458-467, April 1969.
- Martin, David W., Satellite Studies of Cyclonic Developments over the Southern Ocean, Ph. D. Thesis, University of Wisconsin, Madison, Wis., 1968.
- Pomalaza, J. C., "Remote Sensing of the Atmosphere by Occultation Satellites," Stanford Electronics Laboratory, Tech. Report 3670-1, Sept. 1969.
- Sargeant, D. H., "Ray Analysis of the Retroaction of Microwaves Propagated Between Satellites in a Spherically Stratified Medium," Studies of Atmospheric Energetics Based on Aerospace Probings, Annual Report—1967, Dept. of Meteorology, U. of Wisconsin, pp. 61-104, May 1968.
- Smith, William L., H. M. Woolf and W. C. Jacob, A Regression Method for Obtaining Real Time Temperature and Geopotential Height Profiles from Satellite Spectrometer Measurements and its Application to Nimbus III "SIRS" Observations, submitted for publication to Monthly Wea. Rev., 1970.
- Stanford University, SPINMAP, Final Report, June 1966.
- Thayer, G. D., "A Rapid and Accurate Ray Tracing Algorithm for a Horizontally Stratified Atmosphere," Radio Science, Vol. I (New Series), 249-252, Feb. 1967.

N71-11624

BALLOON-BORNE RADIO ALTIMETER

Nadav Levanon
Student Member, IEEE

ABSTRACT

This work describes an accurate pulse radar altimeter which is simple and light enough to be carried aloft by a regular sounding balloon. The altimeter uses a single superregenerative stage serving as both the receiver and the transmitter. This stage is used in a feedback system in such a way that the period between transmitted pulses is a measure of altitude. The relatively slow rate of change of altitude allows the averaging of many returns with corresponding improvement in signal-to-noise ratio.

Accurate geometric altitude of a meteorological balloon can improve atmospheric sounding. It can also replace either temperature or pressure measurements through the use of the hydrostatic equation. Geometric altitude is essential in floating superpressure balloons, where, together with a pressure measurement, it yields the necessary reference pressure.

Analysis of the altimeter and details of the instrumentation are given, along with results of flight tests. Altitudes up to 20 km were measured in balloon flight tests over Lake Michigan. At midrange, the rms value of the random error was smaller than 0.07 percent. Height readings were available every 2 seconds during both the ascent and descent of the balloon payload. The peak power of the altimeter was one watt.

Introduction

The radiosonde, an instrument package elevated to high altitudes in the atmosphere by a buoyant balloon, is the most widespread means of obtaining vertical profiles of atmospheric parameters. These parameters are: temperature,

This research was supported by NASA under Contract NAS5-11542.

The author is with the Space Science and Engineering Center, The University of Wisconsin, Madison, Wis., 53706.

pressure, relative humidity, and wind velocity. The meteorologist is particularly interested in the change in these parameters with altitude. Sensors for the first three parameters are mounted in the balloon package, and their readings are transmitted to the ground. The wind is studied by tracking the balloon with optical or radio theodolites. The only parameter that is not measured directly is the altitude. However, we can allow one unmeasured parameter, because for thin layers of the atmosphere below 90 km, the following relation exists:

$$\frac{P}{P_b} = \exp\left(-\frac{g_0 M \Delta h^2}{R(T + T_b)}\right) \quad (1)$$

where

- P = the pressure at the top of the layer
- P_b = the pressure at the base of the layer
- T = the temperature at the top of the layer
- T_b = the temperature at the base of the layer
- g₀ = acceleration due to gravity, at sea level
- M = molecular weight of the air
- R = universal gas constant
- Δh = the thickness of the layer (in geopotential meters).

From (1) we see that once the pressure and temperature at a certain altitude (usually the surface) are known, the height measurement is not needed to get the vertical profile, as long as the temperature and pressure are measured at close intervals. However, at high altitudes these parameters are hard to measure accurately. For example, the barometer readings are in steps of 1 mbar, which is 0.1 percent of the total pressure at the surface, but which is 10 percent at an altitude of 30 km. The temperature readings are also less accurate at high altitude as a result of stronger solar radiation. A direct measurement of altitude at this range could improve the accuracy of the vertical profile.

The meteorological community is devoting much effort to several programs of global weather observations. Playing major roles in these programs are super-pressure balloons, carrying their sensors at constant-density altitudes and floating with the winds for several months. These balloons will float most of the time over the ocean where no ground stations exist to give the necessary reference readings of surface pressure. For these balloons, a radio altimeter and a pressure measurement could supply such a reference.

The simultaneous reading of pressure and geometric altitude is most useful when we wish to study an unbalanced atmosphere. For example, the balloon in a balanced atmosphere will follow the wind along isobaric lines. Therefore, constant-pressure readings will be accompanied by constant-altitude readings, but when the geometric altitude changes separately from the pressure altitude,

the wind crosses isobars, implying that a weather disturbance is growing or filling.

The above are just a few examples of the great need for a radio altimeter on board the balloon package. However, conventional radio altimeters are too heavy for balloons. It seems that just using modern circuit techniques like integrated circuitry would not be enough in reducing weight, size, and price; and that in addition, a simpler system is needed. It should be noted that the balloon offers one big advantage—its low rate of altitude changes (compared to rockets, for example). This allows plenty of time for averaging, which in turn can replace strong signals, or a good noise figure. This advantage is fully exploited in the proposed balloon-borne radio altimeter.

Principle of Operation

The proposed radio altimeter uses a single superregenerative stage [1] serving as both the receiver and the transmitter. This method was used for beacons in the early days of radar. This superregenerative stage is used in a feedback system in such a way that the period between transmitted pulses is a measure of altitude. The following section describes this operation in more detail.

The superregenerative stage, when triggered by the quench pulses, produces RF pulses whose envelope area depends on the RF signal existing at the input to the stage at the beginning of each quench pulse. The operation of a superregenerative stage is described in more detail in the next section.

With no signal, the noise is the only input. When there is a signal, and when the period of the quenching pulses is equal to the delay of the return pulse, the superregenerative stage detects its own pulses (Fig. 1). As we change the period (or repetition rate) of the quench pulses, we get a maximum output from the superregenerative detector near the repetition rate,

$$f_q = \frac{c}{2h} \quad (2)$$

where

- f_q = the repetition rate of the quench pulses
- c = velocity of light
- h = the altitude.

Varying the repetition rate around this value is equivalent to scanning the returned pulse. As we will see later, the returned pulse is close to an imperfect integration of the transmitted pulse, and, therefore, has a definite peak at the end of the transmitted pulse.

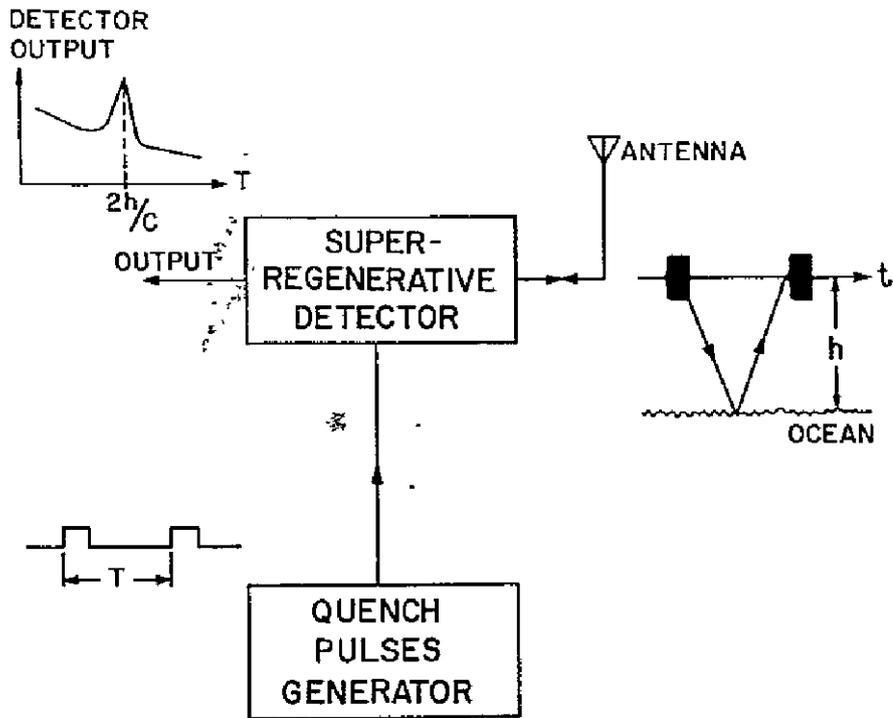


Fig. 1. The output of the superregenerative detector peaks when the period of the quench pulses equals the delay of the returned pulse.

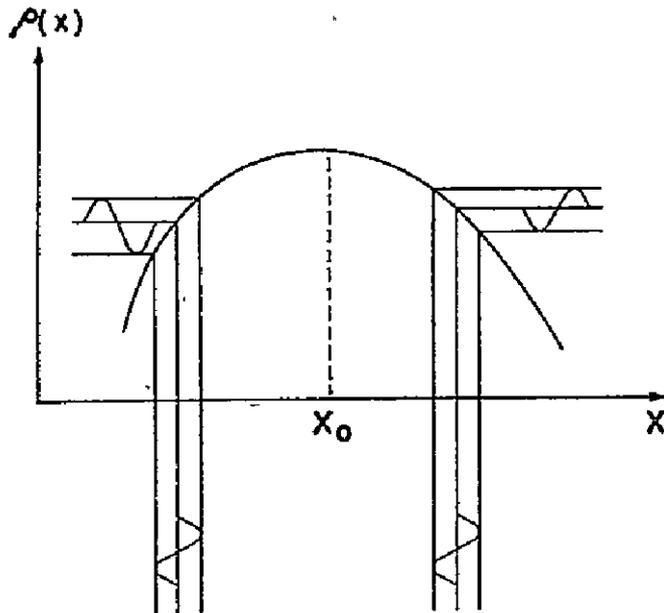


Fig. 2. The effect of sinusoidal parameter perturbations upon a single-dimensional even-function curve.

The rest of the circuitry is a servo loop designed to lock on that peak. This is an example of a "single-dimensional sinusoidal perturbation adaptive system," described in detail in the current literature [2].

The principle of such a system is shown in Fig. 2. This figure demonstrates how a sinusoidal perturbation of the parameter X , on which we have control, serves to take the derivative of a given function, $\rho(X)$, of this parameter. At the peak, the output signal of the perturbation frequency is zero, while off the peak there will be some slope, either positive or negative, and therefore an output signal, in phase or out of phase, respectively. This type of signal can be used as an error signal in a servo loop.

The block diagram of the system is shown in Fig. 3. The voltage controlled oscillator (VCO) determines the repetition rate of the quench pulses. This rate is linearly related to the voltage at the output of the integrator and is sinusoidally perturbed. In other words, the quench frequency is frequency modulated at the perturbation frequency f_p ; the deviation of this FM is very small. The curve of the output of the superregenerative detector as a function of the quench frequency serves as an FM detector for the perturbation signal. When the

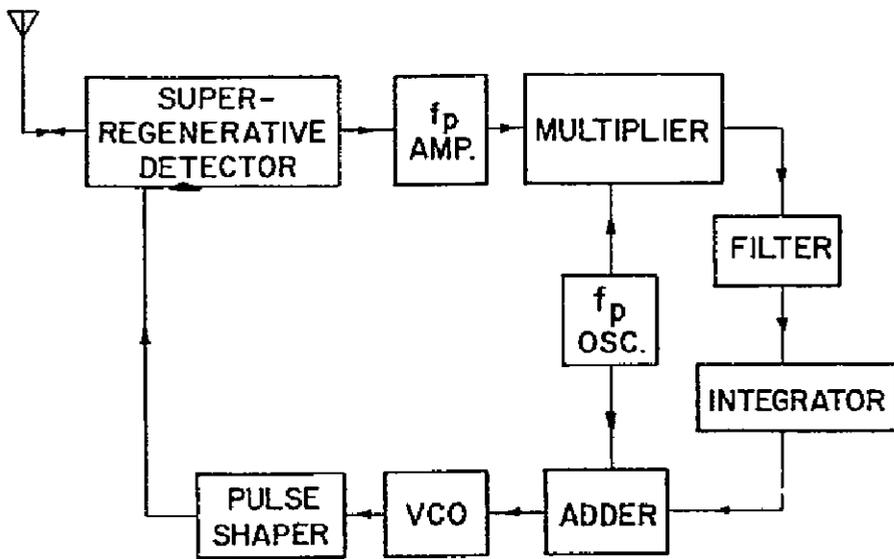


Fig. 3. Block diagram of the altimeter.

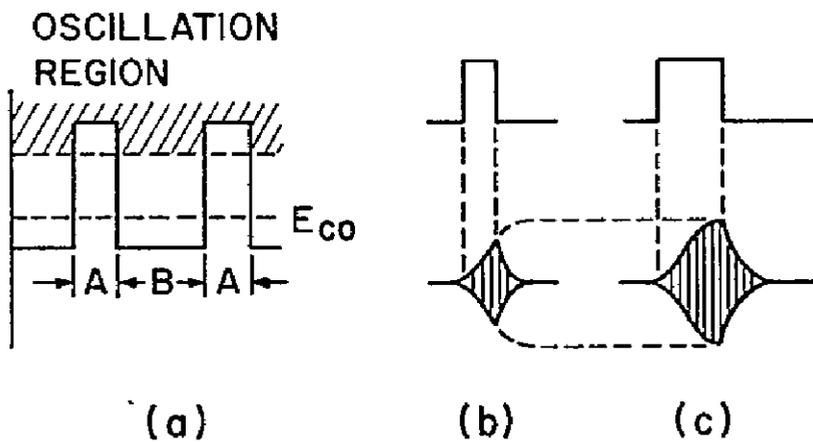


Fig. 4. Quench voltage (a) and relation to modes of oscillation (b) and (c).

quench frequency is too high, the detected perturbation signal is out of phase, compared to the modulating perturbation signal; when f_q is too low, the perturbation signal is in phase; and when it is at the right frequency, we get only the second harmonics of the perturbation signal.

The output of the detector is passed through a narrow-band amplifier centered around f_p and then multiplied by the modulating signal. The output of this multiplier contains a positive dc voltage when the two signals are in phase, and a negative dc voltage when they are out of phase. The integrator output is affected by this dc voltage, and locks the VCO to that frequency which gives the peak at the output of the superregenerative detector.

The Superregenerative Detector

The fundamental action in a superregenerative circuit is centered around the growth of oscillations in an oscillator. The variation in grid voltage is shown in Fig. 4(a). Between positive pulses, the grid is at the negative potential which biases the tube beyond E_{CO} , the cutoff voltage. During the positive pulses, the grid potential is raised to a point where the circuit oscillates, as indicated by projection into the shaded oscillation region. Thus, oscillation grows during the A periods and decays during the B periods.

There are two modes of operation defined for superregeneration: a linear mode (Fig. 4(b)), and a logarithmic mode (Fig. 4(c)). The linear mode results when the positive quench period A is so short that the oscillations do not have time to build to full saturation amplitude. The logarithmic mode occurs when the A period is sufficiently long to allow oscillations to build to full amplitude before the end of the A period. In our case the circuit operates also as the transmitter, and we need as much output power as possible. Therefore, we operate in the logarithmic mode.

Before saturation is reached, the envelope of the oscillations will rise as:

$$e_c = V_0 e^{Rt/2L} \quad (3)$$

where

- V_0 = the input voltage when period A begins
- R = the total negative resistance of the circuit
- L = the circuit inductance.

When there is no input signal, the input is the noise voltage V_n . When a signal voltage V_s is present, the voltage at the start of the oscillations will be $V_s + V_n$ and (3) becomes

$$e_c = (V_s + V_n) e^{Rt/2L} \quad (4)$$

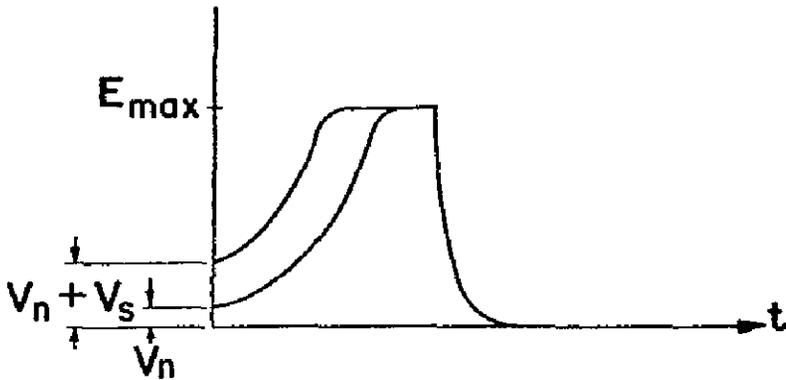


Fig. 5. The change in pulse area in the presence of signal.

As can be seen in Fig. 5, a higher initial voltage would cause a shorter rise time period before saturation, and, therefore, a larger pulse area. The additional area in the signal-originated pulse is added with each additional pulse, and the total change appears at the output of the detector. The change in average voltage of the detected output ΔV , assuming linear detection, is given by [1]

$$\Delta V = f_q E_{\max} \frac{2L}{R} \ln \left(\frac{V_s + V_n}{V_n} \right). \quad (5)$$

This change in output divided by the change in the input gives the gain of the stage. Assume, for example, that the signal is equal to the noise; then the change in the input is equal to the noise, which is on the order of a few microvolts. The change in the output, however, will be

$$\Delta V = f_q E_{\max} \frac{2L}{R} \ln 2, \quad (6)$$

which is of the order of volts. This means that a gain of a million can be achieved in a single stage.

It is for both the high receiver gain and the doubling of the single stage as a transmitter and receiver that the superregenerative stage has been chosen for a small, light, and accurate balloon-borne radio altimeter.

The Returned Pulse

The ocean surface plays the role of an imperfect integrator on the returned pulse. Let us assume, for a moment, that the three following conditions exist:

- 1) The average scattering cross section of the ocean is independent of the incident angle.
- 2) The altimeter antenna is isotropic.
- 3) The additional attenuation due to a longer path length of a side reflection compared to a downward reflection can be neglected.

With these three assumptions, we will follow the pulse on its way to the ocean surface and back. The first reflection occurs when the front end of the pulse reaches the point B (Fig. 6) just below the altimeter. Side returns like that from point C will start later, as a result of the additional distance, and will reach the altimeter even later because the returning pulse also has the additional distance to traverse. While the front edge of the pulse is being reflected from point C, later portions of the pulse reach point B, and start to be reflected.

In other words, the pulse can be seen as traveling on the ocean surface. When the trailing edge is being reflected from point B, the whole pulse is reflected to the altimeter, and the intensity of the return reaches a maximum. With the above three assumptions, this level will remain constant as the pulse on the ground travels away from the altimeter subpoint.

Measurements have demonstrated [3] that with occasional exceptions, a radar return from the ground is largely due to area scatter. For this reason, no phase relations were introduced, and when we speak of the returned pulse, we are referring to the mean pulse envelope.

As we mentioned above, the maximum of the returned pulse occurs at the end of the transmitted pulse (if we ignore the basic delay). If we now remove the assumption that the antenna is isotropic, and, instead, assume that it has constant gain for $|\theta| < \theta_0$ and zero gain outside, we can calculate the minimum θ_0 that still allows the peak to occur at the end of the transmitted pulse.

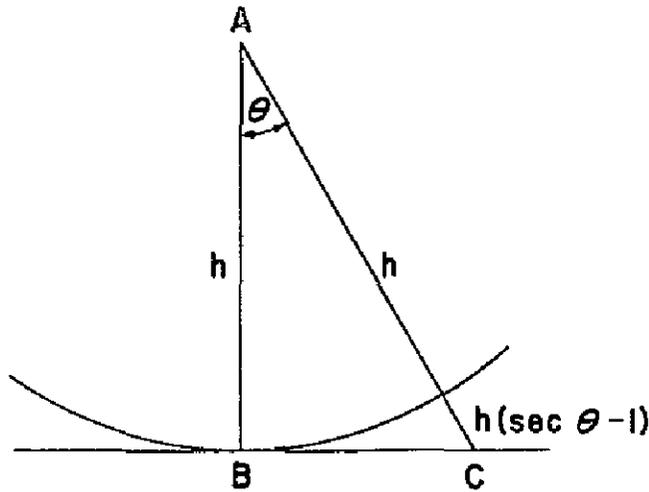


Fig. 6. The geometry of reflection.

From Fig. 6 we see that the return from point C has additional delay τ , which is, if we assume a flat surface,

$$\tau = \frac{2h}{c} (\sec \theta_0 - 1) \quad (7)$$

where

- h = the altitude
- c = the velocity of light.

If the length of the transmitted pulse is shorter than τ , then the peak of the return pulse occurs at the end of the transmitted pulse, which does not change with altitude. This is called "pulse-length-limited illumination." If the length of the transmitted pulse is larger than τ , the peak will occur at τ which in turn is dependent on altitude. This is called "beam-width-limited illumination."

Our system locks on the peak of the returned pulse, and we would prefer the additional delay to be constant; therefore, we use pulse-length-limited illumination. The condition for such illumination is that

$$\sec \theta_0 \geq \frac{c\tau_0}{2h_{\min}} + 1 \quad (8)$$

where

- τ_0 = the length of the pulse
- h_{\min} = the minimum altitude.

When we include the dependence of the average scattering cross section of the ocean on the angle of incidence, the additional attenuation due to additional distance, and the fact that the antenna pattern is not constant as a function of θ , the returned signal can be expressed as the convolution of the transmitted pulse waveform in power units with a function which includes effects of antenna pattern, ground properties, and distance.

Moore et al. [4] shows that for a flat earth and a system that is independent of azimuth orientation, the mean received pulse envelope in power units is given by

$$P_R(d) = \frac{\lambda^2}{8(2\pi)^2} \int_h^{cd/2} \frac{1}{r^3} P_T(d - \frac{2r}{c}) G^2(\theta) \sigma_o(\theta) dr, \quad (9)$$

where

r = the distance from the altimeter to the points of reflection
($r = h \sec \theta$)

h = the altitude

θ = the angle from the vertical

d = the time from the leading edge of the transmitted pulse ($d > 2h/c$)

λ = the wavelength of transmitted energy

$G(\theta)$ = the antenna gain

$\sigma_o(\theta)$ = the average scattering cross section of the ocean

$P_T(\)$ = the envelope of the transmitted pulse in power units .

$P_R(\)$ = the envelope of the received pulse in power units

c = velocity of light

We will include both the antenna pattern and the ground scattering pattern in one equation,

$$G^2(\theta) \sigma_o(\theta) = G^2 \sigma_o \cos^6 \theta = G^2 \sigma_o \left(\frac{h}{r}\right)^6 \quad (10)$$

which is a good approximation in many cases.

Inserting (10) in (9), we get

7

$$P_R(d) \approx \frac{\lambda^2 G^2 \sigma_0 h^6}{8(2\pi)^2} \int_h^{cd/2} \frac{1}{r^9} P_T(d - \frac{2r}{c}) dr. \quad (11)$$

Further approximations lead to

$$P_R(t + \frac{2h}{c}) \approx \frac{\lambda^2 G^2 \sigma_0 c}{16(2\pi)^2 h^3} \int_0^t P_T(t - \tau) \cdot (1 - \frac{9c}{2h} \tau) d\tau \quad (12)$$

where t is the time from the reception of the leading edge.

In Fig. 7 a typical superregenerative pulse is used as the transmitted pulse, P_T . When this envelope is used in (12), the envelope of the received pulse is as shown in Fig. 7.

For pulses of length in the order of $1 \mu s$, the convolution integral can be approximated by a regular integral, and this will give a simpler expression for the peak of the received pulse P_2 :

$$P_2 = \frac{\lambda^2 G^2 \sigma_0 c}{16(2\pi)^2 h^3} P_{\max}(T - t_1). \quad (13)$$

The only constant in (13) that requires clarification is the average scattering cross section σ_0 . Edison *et al.* [3] measured this value for vertical incidence at 415 MHz and found it to vary from 0.7 for dense woods, to 4 for certain city targets, and up to 50 for slightly rough water.

Differences between water and land exist also in the range of fading (defined here as the range between the level below which only 5 percent of the return power is found, and the level below which 95 percent is found). This range was as small as 3 dB for a smooth water surface and as large as 19 dB for certain nonhomogeneous target areas. The effect of this fading range could not be included without knowing the correlation of the signal fluctuations and the number of pulses averaged. The last question can be answered immediately. If the altimeter averaging period is about one second and the smallest repetition rate is 5×10^4 pps, then at least 50,000 pulses are averaged. As to the correlation of the signal fluctuations, data was unavailable for such a high repetition rate. Edison *et al.* [3] showed that at a repetition rate of 400 pps, over wooded area, consecutive returns do not resemble each other at all. From this we can conclude that at this rate the correlation coefficient is close to zero.

For the case of uncorrelated signal fluctuations, Schwartz [5] has shown that if the number of pulses averaged is larger than about 100, and the instantaneous reflected power obeys a Rayleigh probability-density function, which is typical for large targets, then the fluctuations have no effect on the

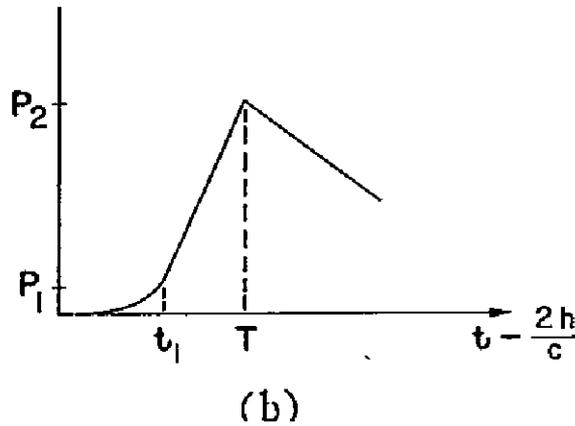
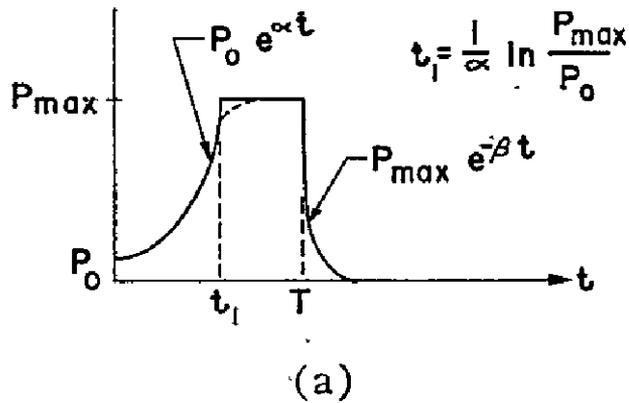


Fig: 7. The transmitted (a) and received (b) pulses.

detectability of the signal. Accepting Edison's results we know that we average at least 400 uncorrelated pulses. We can therefore assume that the mean returned pulse is a good approximation of the returned pulses.

Let us introduce some typical values in (13):

$$P_{\max} = 1 \text{ watt}$$

$$P_0 = 4 \times 10^{-12} \text{ watt}$$

$$T - t_1 = 0.5 \times 10^{-6} \text{ second}$$

$$\sigma_0 = 1 \text{ (for land), } 20 \text{ (for water)}$$

$$G = 3.5 \text{ (Yagi), } 13 \text{ (rhombic)}$$

$$\lambda = 0.7 \text{ meter,}$$

TABLE I
THE EXPECTED RANGE OF THE ALTIMETER

Surface Antenna	Land $\sigma_0 = 1$	Water $\sigma_0 = 1$
Yagi $G = 3.5$	7 km	20 km
Rhombic $G = 13$	17 km	46 km

and ask what is the height h at which $P_2 = P_0$. The answer is summarized in Table I.

Linear Analysis of the Altimeter

The sinusoidal perturbation serves to take the derivative of a curve which is actually the dependence of the superregenerative detector output on the input to the VCO. This curve includes in it the delay, intensity, and shape of the returned pulse, the linear voltage to frequency characteristic of the VCO, and the logarithmic gain of the superregenerative detector. To simplify our analysis we will assume this curve to be a parabola. The amplitude of the signal in the perturbing frequency, f_p , as it appears after the amplifier is determined by the derivative of this parabola, and is therefore a linear function of the distance from the peak of the parabola.

We are thus able to draw a linear simplified equivalent block diagram of the altimeter Fig. 8. In this block diagram c_1 is the amplitude of the perturbing signal, ρ'' is the second derivative of the parabola (a constant). A is the total gain, and it includes the gain or attenuation of

- 1) the RF path
- 2) the superregenerative circuit
- 3) the bandpass amplifier
- 4) the integrator ($\approx 1/RC$)
- 5) the VCO
- 6) the multiplier ($\approx 1/2$).

$F(s)$ is the Laplace transform of the total response of the networks in the loop, and it includes

- 1) the bandpass response of the f_p amplifier transferred to low pass
- 2) the filter response.

Of these two, the second is the dominant.

$1/s$ represents the integrator, and $n(t)$ is the noise voltage; $n(t)$ is multiplied by $1/\sqrt{2}$ to indicate that the synchronous detection of the perturbation signal eliminates half of the noise power accompanying it. The output of this loop could be either in voltage form, if it is measured at the input to the VCO, or in frequency form, if it is measured at the output of the VCO. This frequency, which is equal to the pulse repetition rate, is in essence the height information.

The loop of Fig. 8 resembles a typical phase-locked loop. We can thus use studies made on this type of loop [6], [7], in order to optimize the output error due to noise, the steady-state error, and the acquisition behavior of the loop.

The conclusions drawn from phase-locked loops led to the inclusion of a lag-lead network as the loop filter, to a choice of damping factor ξ of 0.7, and loop-noise bandwidth B_L of 1 Hz.

The primary noise source of the altimeter is the superregenerative stage. As Hall [1] showed, the effect of the sampling done in the relatively broad RF stage is to multiply the effective noise bandwidth of the rest of the receiver by twice the ratio between the RF bandwidth Δf_0 and the sampling rate f_q .

Viterbi [6] has shown that in a linear phase-locked loop the mean square error due to noise σ_x^2 is equal to the inverse of the signal-to-noise ratio at the output, $(SNR)_0$, i. e.,

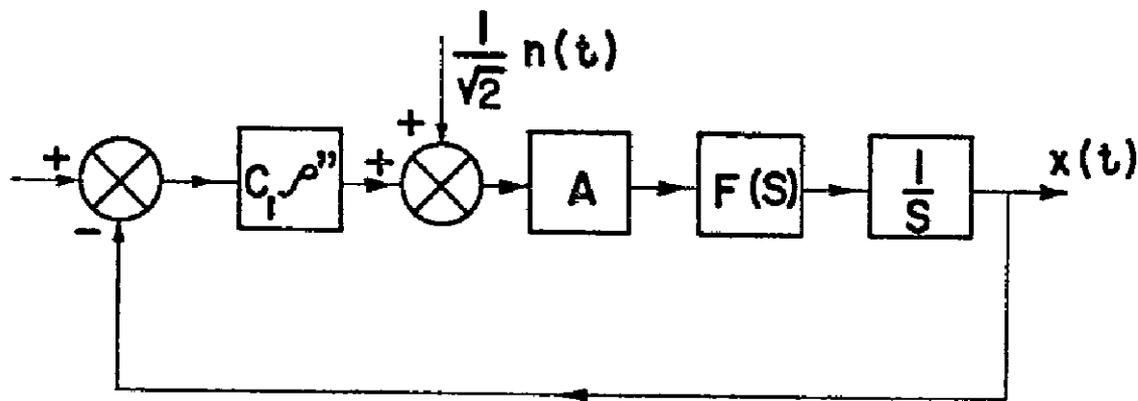


Fig. 8. A simplified equivalent block diagram of the altimeter.

$$(\text{SNR})_0 = \frac{1}{\sigma_x^2} \quad (14)$$

We can now give a crude yet meaningful answer to the following question. Given the allowed mean square error σ_x^2 , what is the minimum signal-to-noise ratio per each pulse at the input of the altimeter?

To answer this question we shall use Fig. 9 in which the various processes which the signal passes, and their effect on the signal-to-noise ratio, are shown.

The first term is the $1/F$ change in signal-to-noise ratio due to the noise figure F of the RF stage. The second is the increase in noise due to sampling. The \sqrt{n} improvement is due to averaging of n pulses [8].

The term 2 is a result of the synchronous detection, and the last term is the improvement in signal-to-noise ratio due to the change from the RF bandwidth Δf_0 to the loop bandwidth B_L .

We note that the number of pulses averaged, n , is given by

$$n = \frac{f_q}{B_L} \quad (15)$$

where

f_q = the repetition rate

B_L = the loop effective bandwidth.

Thus we get from Fig. 9 that the signal-to-noise ratio at the output, $(\text{SNR})_0$, is related to the signal-to-noise ratio at the input, $(\text{SNR})_i$, as follows:

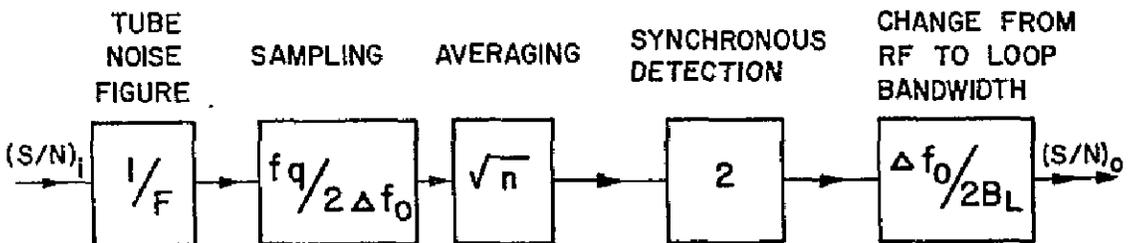


Fig. 9. Signal-to-noise ratio budget.

$$(\text{SNR})_o = \frac{1}{2F} \left(\frac{f_q}{B_L} \right)^{3/2} (\text{SNR})_i \quad (16)$$

Using (14) we get

$$(\text{SNR})_i = 2F\sigma_x^{-2} \left(\frac{f_q}{B_L} \right)^{-3/2} \quad (17)$$

Equation (17) is the answer to the question we have just asked. However, we will go one step further to make it more meaningful.

The noise power at the input is given by

$$N_i = K T \Delta f_0, \quad (18)$$

where

K = Boltzmann's constant

T = noise-source temperature, °K.

Inserting (18) in (17) we get for the signal power input

$$S_i = K T \Delta f_0 2F \sigma_x^{-2} \left(\frac{f_q}{B_L} \right)^{-3/2} \quad (19)$$

It will be interesting to put numbers in this equation. We recall that σ_x^2 is the mean square error due to noise. If the total error allowed is 0.03 percent, a sensible choice of σ_x will be 0.01 percent or 10^{-4} ; thus, we have for the balloon flight model:

$$KT = 5 \times 10^{-21} \text{ W/Hz}$$

$$\Delta f_0 = 5 \times 10^6 \text{ Hz}$$

$$F = 8$$

$$\sigma_x = 10^{-4}$$

$$f_q = 10^5 \text{ Hz}$$

$$B_L = 1 \text{ Hz}$$

and the resulting signal power required at the input is $S_i = 10^{-12}$ watts, which is 7 μV over 50 ohms.

The mean square error due to noise could be further lowered if a smaller loop bandwidth B_L were chosen. The theoretical lower limit on B_L should have been determined by the rate of change of height that the altimeter has to follow. However, some practical considerations enter here. The total delay range that the altimeter can measure is considerably larger than the deviation of the returned pulse. Unless the VCO period falls within the delay range covered by the returned pulse, the loop is not closed and no acquisition is possible. To overcome this problem, a constant voltage is fed to the integrator to cause a certain sweeping rate* of the VCO frequency, and thus, the whole range is scanned. This sweeping rate causes a steady-state error; and in addition, it could carry the transient error beyond the hold-in range, and thus prevent locking altogether. The locking behavior of the loop in the presence of this sweeping rate is similar to phase-locked loop tracking a frequency step. For the linear model it was found [9] that the relation between the sweeping rate and the lower limit to the loop-noise bandwidth B_{Lmin} is given by

$$B_{Lmin} = \frac{\xi + \frac{1}{4\xi}}{2\Delta T(0.6 + 2\xi)} \quad (20)$$

where ΔT is the time (in seconds) required to scan the leading edge of the returned pulse. In laboratory tests of the altimeter it was found that the practical lower limit is about three times larger than the theoretical.

Again, we shall insert practical numbers in (20). The delay range covered by the altimeter is 12 to 27 μ s, and it is scanned in 30 seconds; the leading edge of the returned pulse covers 0.5 μ s. From this we can conclude that $\Delta T = 1$ second. With the damping factor $\xi = 0.7$, we get for the theoretical lower limit on the loop-noise bandwidth:

$$B_{Lmin} = 0.27 \text{ Hz.}$$

The practical lower limit should be 3 times larger, and this would lead us to the choice of a loop-noise bandwidth of 1 Hz.

Ambiguity

There is no reason why the altimeter cannot transmit a second pulse before the previous one is received. We shall call this subharmonic operation (Fig. 10). This mode of operation can happen by itself at any altitude above twice

* It should be pointed out that some sweeping rate is inherent in the system, and some is built-in due to circuitry offsetting.

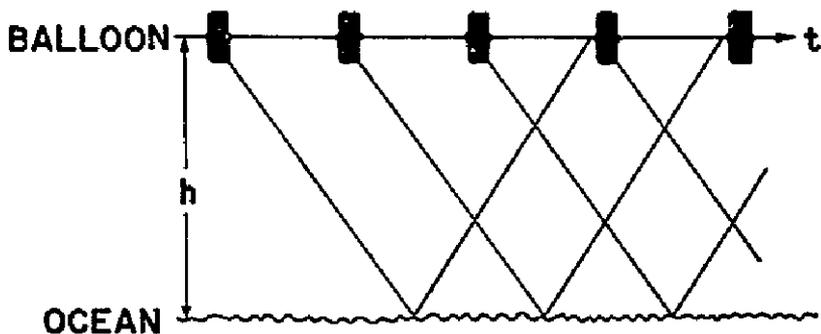


Fig. 10. Subharmonic operation

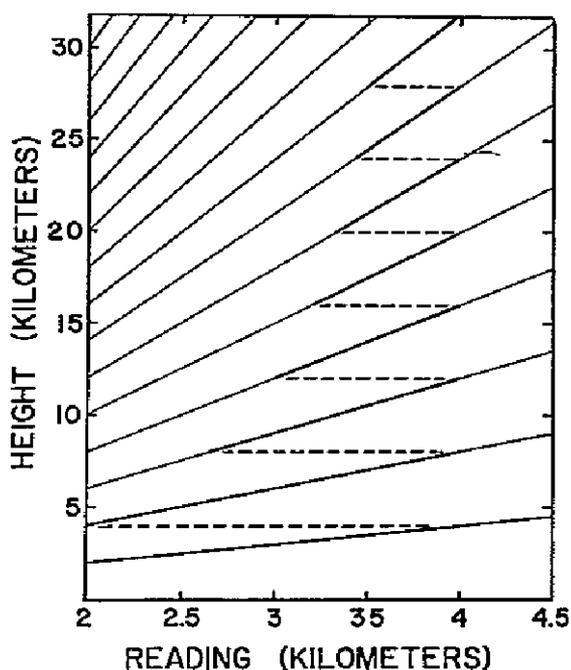


Fig. 11. Ambiguity pattern.

the minimum altitude of operation. This immediately indicates that the minimum range of altitudes required to get full coverage is equal to the minimum altitude. If the minimum altitude of operation is 2000 meters, then the range of the altimeter should be from 2000 meters to 4000 meters. Above 4000 meters the altimeter will switch to a subharmonic mode. From Fig. 11 it is clear that the first ambiguity problem starts at 6000 meters, which is 3×2000 meters, or 2×3000 meters, and the altimeter can lock on both. The pattern of switching modes as the balloon ascends, in the case of no loss of locking except at the end of range, is marked in Fig. 11. However, this pattern is not guaranteed since loss of locking might occur. In any case, the ambiguity is at least the lowest altitude (in our example 2000 meters), and it can be resolved by pressure reading or by the history of the ascending.

Subharmonic operation is recommended even if the range of operation is limited, as in the case of the superpressure balloons. In this case, however, it is possible to assure that only one mode will cover this limited range. Operation with subharmonics means higher quench frequency f_q and therefore better signal-to-noise ratio.

Circuitry, Laboratory Test System, and Ground Station

This section briefly covers the main features of the altimeter circuitry and the methods used for laboratory testing and reception of altimeter data.

The superregenerative stage utilizes a nuvistor triode for the 403 MHz version, and a pencil tube with integral resonator (the type used in radiosondes) for the 1680 MHz version. The narrow-band amplifier is an operational amplifier with a twin T network in the feedback loop. The multiplier is a ring demodulator. The integrator and oscillator utilize operational amplifiers. The VCO is an astable multivibrator in which the charging resistors are returned to a variable input. The pulse shaper includes a monostable stage that determines the duration of the quench pulse and an amplifier.

In addition, a bias control stage which is fed by all the supply voltages involved in the superregeneration action keeps the tube bias at the correct level in order to assure proper superregeneration. Altogether there are seven integrated circuits, one transistor, and one vacuum tube.

Two types of antennas were used in the flight test of the 403 MHz version. The first was a rhombic antenna with leg length of 3λ , and tilt angle of 58.5° . The measured gain of the rhombic was 9 dB over a half-wave dipole.

The second antenna was a two-element Yagi, with the parasitic element as the director, and a spacing of 0.13λ between the elements. The measured gain of the Yagi was 3.2 dB over a half-wave dipole.

The laboratory test system appears in Fig. 12. The quench pulses were used to trigger a delayed pulse generator. These delayed pulses modulated a UHF signal generator whose output simulated the returned pulses. The output of the signal generator was connected to the input (and also output) of the superregenerative detector. This point was also connected to a detector in front of an oscilloscope. This scope was used to monitor the shape of the envelope of the transmitted pulses. The counter was used to count the repetition rate. The chart recorder measured the input voltage to the VCO. These two pieces of equipment were used to measure steady-state error and noise error. The chart recorder was also used to measure the loop bandwidth. To do this the delay was modulated by a very low-frequency (VLF) sine wave from the function generator. The dual chart recorder recorded both the sinusoidal signal at the output of the function generator, and the resulting sinusoidal

signal at the output of the integrator. By comparing these two amplitudes at several frequencies it was possible to draw the frequency response of the loop.

The amplitude of the simulated returned pulse was controlled by the output attenuator of the signal generator. It should be noted that changes in the returned pulse amplitude affected mainly the signal-to-noise ratio but also the gain of the loop, as the gain includes it in the second derivative of the envelope of the returned pulse.

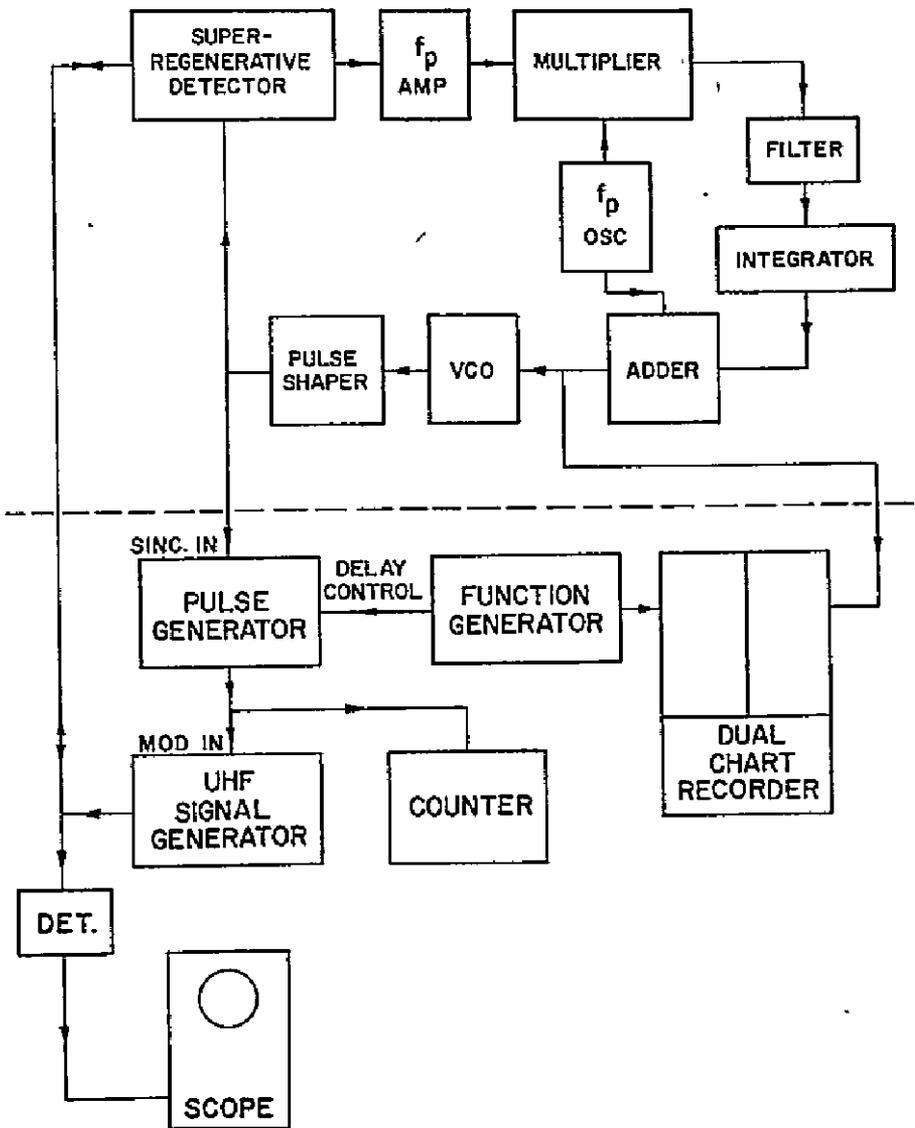


Fig. 12. Block diagram of the altimeter and the test system.

Without any sweeping voltage, the altimeter starts to lock when the peak of the returned pulses is as low as $2.5 \mu\text{V}$. With proper sweeping voltage (the entire range is scanned in 25 seconds), good locking (with mean time to unlock on the order of hours) is achieved with $8 \mu\text{V}$ input. However, the steady-state error is about 0.3 percent and the noise error 0.1 percent. With an input of $15 \mu\text{V}$ the total error is below the specified 0.03 percent.

Data reception on the ground is accomplished by receiving and counting the pulse repetition rate. Two methods are used; either 1) receiving the altimeter RF pulses directly; or 2) modulating the VCO, or quench pulses, on the radiosonde signal and receiving those. In both cases the receiver at the ground station was to have IF and video bandwidth capabilities equal to the highest repetition rate, which, in our case is about 100 kHz. The video output of the ground station receiver is thus a carrier varying slowly between 50 to 100 kHz. From the altimeter loop bandwidth we can deduce that the information bandwidth is less than one hertz. This type of signal naturally requires a phase-lock demodulator for best reception. Yet, despite the fact that the information bandwidth is 1 Hz, the phase-lock loop bandwidth has to be at least f_p , which is the perturbation frequency. The block diagram of the ground station appears in Fig. 13.

Flight Tests

As of March 31, 1969, a total of 4 flight tests were conducted. The purpose of the first one was to test the principle of operation, i. e., whether the altimeter would lock to the returned pulse, and how it would follow changes in altitude. This test was conducted on June 5, 1968, over Lake Michigan, from an aircraft provided by the National Center for Atmospheric Research, Boulder, Colo. The antenna used in the flight was a slotted dipole. Excellent locking was achieved up to the maximum altitude of the flight—8000 feet. During the 30-minute period over the lake, locking was not lost even once, unless by external disturbance by the operator. Readings of altitude were performed by reading the period of the quench pulses with a portable oscilloscope. These

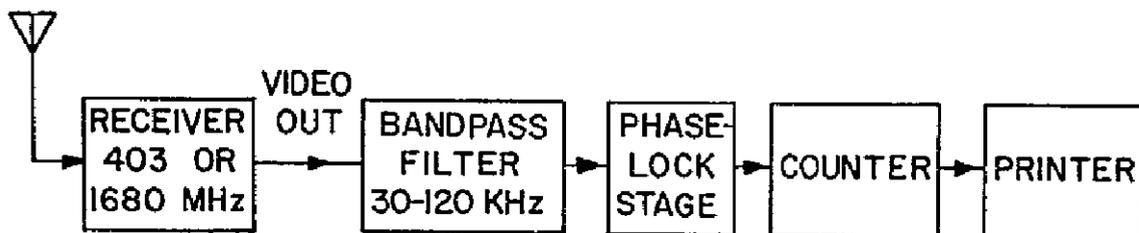


Fig. 13. Block diagram of the ground station.

readings were compared to readings of the pressure altimeter of the aircraft. With these rather limited facilities, readings on the tested altimeter were within ± 0.1 percent of the readings on the aircraft altimeter.

The height of the waves in Lake Michigan at the time of the above test was 3 feet according to the U. S. Coast Guard. En route back to Madison, the altimeter was still locking over land at an altitude of 8000 feet. However, locking was lost from time to time.

In the second flight test, the radio altimeter, together with a 1680 MHz radiosonde, was carried aloft by a standard 1200 gram neoprene balloon. The altimeter was launched from Madison, Wis., on November 22, 1968, and landed on a farm in Kenosha County, Wis. The rhombic antenna design was used in this flight. Power was drawn partly from the radiosonde battery (100 V and 6 V) and partly from additional water activated batteries (+ 12 V and -12 V). The balloon and the payload of this flight are shown in Fig. 14 a few seconds after being launched. The items from the top are: the balloon, two parachutes, radiosonde, a package with the additional batteries, a bag with a letter to the finder, the radio altimeter, and the rhombic antenna. As far as altitude measurements were concerned, this flight was almost a complete failure. There was only a short intermittent period of about 10 minutes when locking was achieved. This happened at an altitude of 22 km which is rather high for an overland flight. The reason behind this is that, due to poor temperature and voltage compensation, the sweeping rate of the altimeter was decreasing, and at very slow sweeping rate its sensitivity improves considerably. Despite the failure in achieving meaningful altitude readings, many helpful conclusions were drawn from this test. The major one was that the method of counting the altimeter pulses directly from the video signal at the ground station receiver was inadequate as a result of noise. This conclusion led to the addition of a phase-locked stage between the receiver and the counter. Another observation from this test was that the radiosonde signal on 1680 MHz was stronger than the altimeter signal on 403 MHz. This was probably due to: 1) the fact that the altimeter antenna was pointing down while the radiosonde antenna was pointing sideways toward the ground station; 2) the higher frequency of the radiosonde which allows the use of a dish antenna at the ground station, with considerably more gain than the helix antenna used for the 403 MHz signal. This resulted in a change in which the altimeter VCO square wave is used to modulate the radiosonde 1680 MHz transmitter. Thus, the height information is transmitted to the ground through a better channel.

The third and fourth tests were also balloon flight tests. They were launched near Sturgeon Bay, Wis., on March 14 and 15, 1969. The balloon flight test of March 14, 1969, was the most successful thus far, and will be described in detail. In this test the altimeter utilized a 2-element Yagi antenna; the electronics and the antenna are shown in Fig. 15. The electronics weighs 130 grams and the antenna 40 grams. Separate batteries were used in

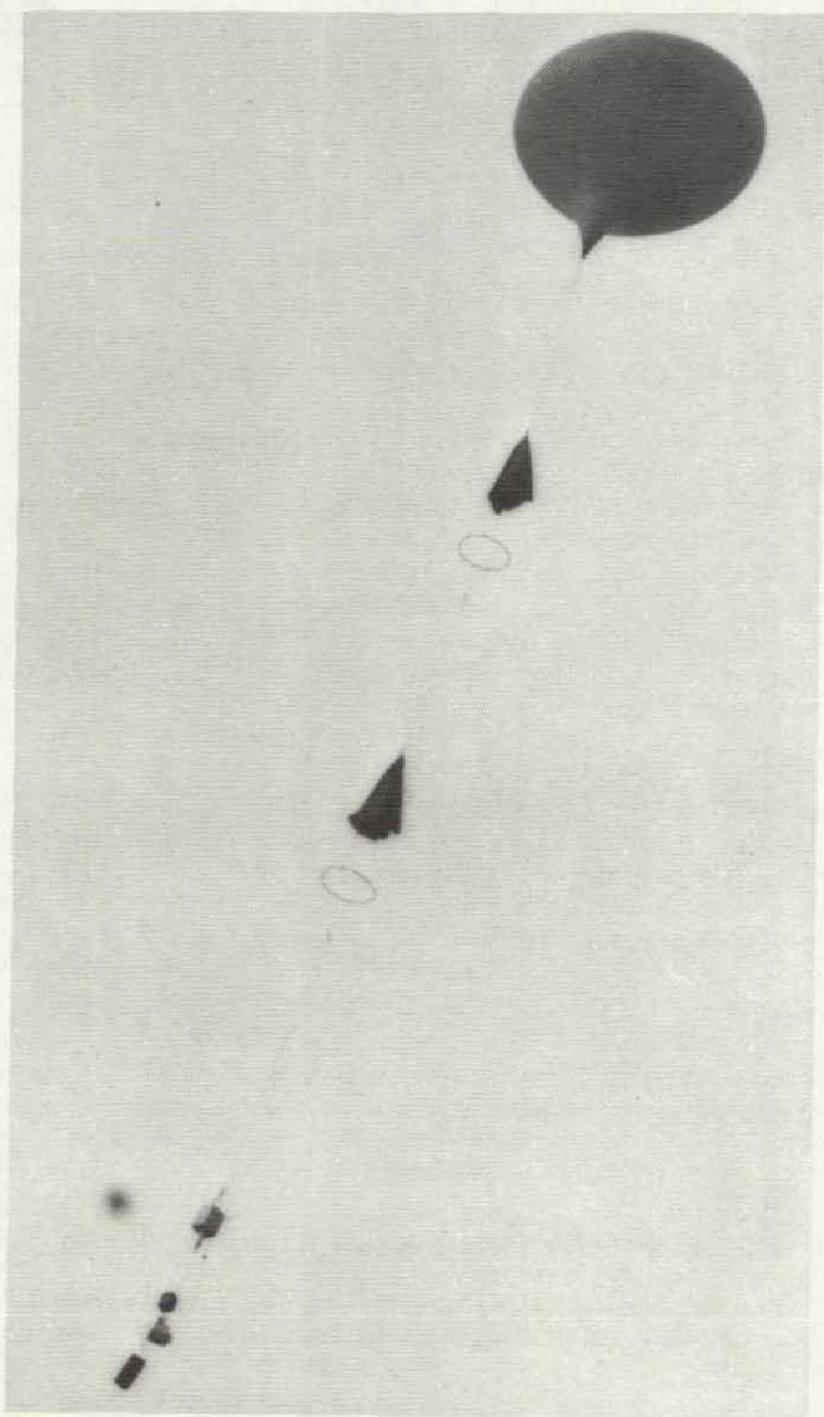


Fig. 14. The balloon payload.

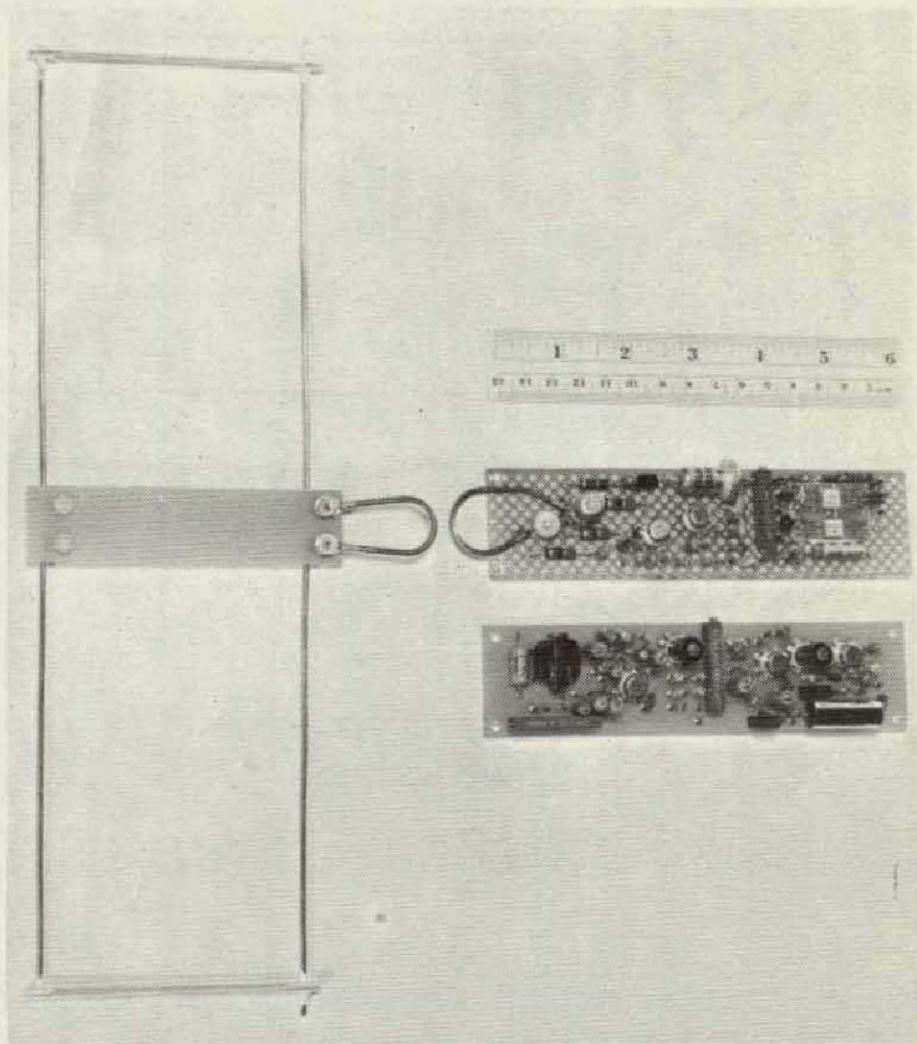


Fig. 15. The radio altimeter layout.

this flight, so that the only electrical connection between the altimeter and the radiosonde is that of the altimeter VCO square wave, which modulated the radiosonde transmitter. Thanks to a calm day, we were able to use a very long string (about 300 feet) between the balloon and the payload, and thus minimize the pendulum motion. After burst of the balloon, the payload landed at Norwalk, Mich., and was returned to us. From that we can conclude that most of the flight, and certainly all of the ascent, took place over Lake Michigan.

Height readings started 7 minutes after launch at an altitude of 1.8 km, and were received for the next 57 minutes at intervals of 2 seconds, until the balloon reached an altitude of 16 km. In the next 10 minutes the altitude readings became intermittent. Yet, good data were still available up to an altitude of 18.5 km, where the signal-to-noise ratio of the radio altimeter decreased to the point that no locking could be achieved.

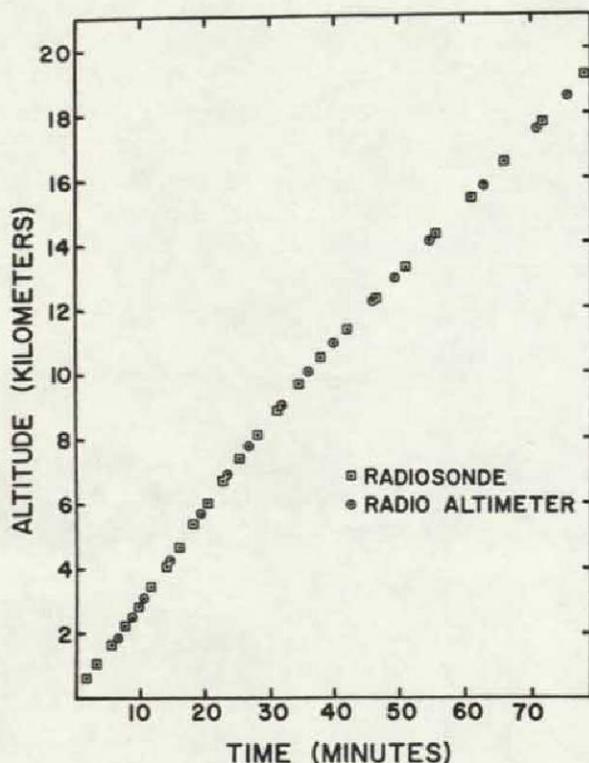


Fig. 16. Altimeter height readings, compared to radiosonde height calculations, during the ascent of March 14, 1969, Sturgeon Bay, Wisconsin.

In Fig. 16 several height measurements of the altimeter in intervals of about 5 minutes, during the ascent, are compared to heights calculated from the temperature and pressure data received from the radiosonde. These calculations utilize the hydrostatic equation (1), where the layers Δh were between every fifth contact on the radiosonde baroswitch. Calibration of the baroswitch was performed two weeks before the flight.

Height readings were received also on the descent, starting at an altitude of 14.2 km and continuing down to 2.2 km, although the readings were intermittent during portions of the descent. Some of these readings are compared to the radiosonde data in Fig. 17.

The altimeter data, which is actually the repetition rate f_q , were recorded in a form of a 5-digit printer output every 2 seconds. The height h is related to the repetition rate through the following equation:

$$h(\text{meters}) = \frac{c(\text{m}/\mu\text{s})}{2} \left[\frac{n \times 10^6}{f_q(\text{Hz})} - \tau(\mu\text{s}) \right] \quad (21)$$

where τ is the width of the quench pulse (which was $1.40 \mu\text{s}$ in this case), c is the velocity of light in meters per μs , and n is the subharmonic number.

Altogether, more than 2500 meaningful height measurements were recorded during the flight. It is impossible to present all the data in this report. We chose the 200 mbar level, which happens to be the midrange of this flight, and presented readings received from the altimeter over a period of two minutes during the ascent (Fig. 18) and one minute during the descent (Fig. 19). From Fig. 18 we note that the 2 seconds rms value of the random error, at an altitude of nearly 10 km, is smaller than 7 meters. This is an error of 0.07 percent.

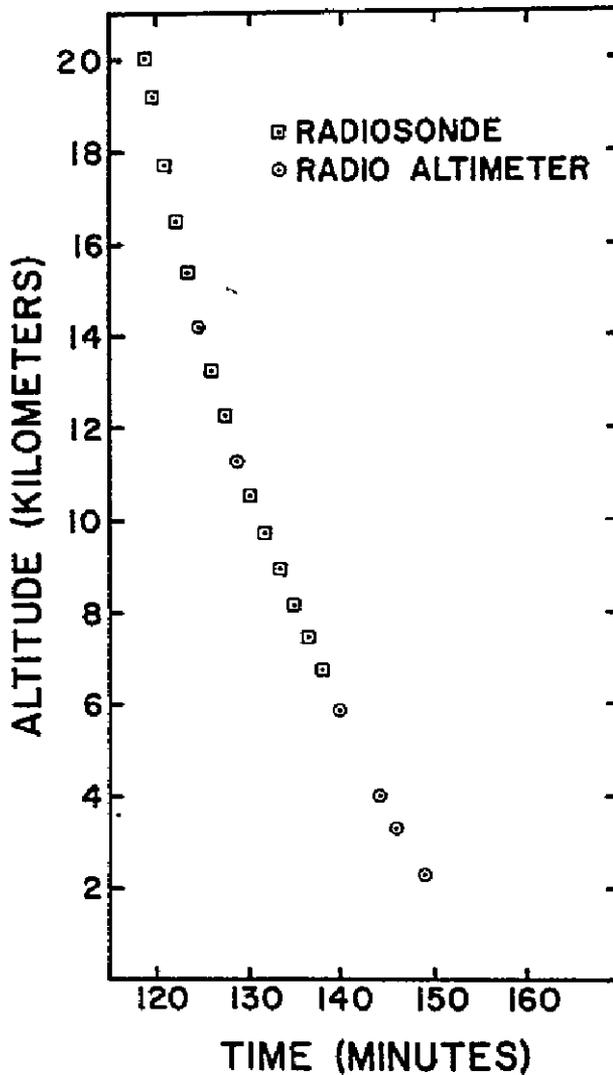


Fig. 17. Altimeter height readings, compared to radiosonde height calculations, during the descent of March 14, 1969, Sturgeon Bay, Wis.

There is no simple way to measure the absolute measurement error of the radio altimeter. The altitude, as calculated from the radiosonde pressure and temperature readings, is certainly less accurate. However, it should be pointed out that the difference between the radiosonde calculated height and the altimeter measured height was smaller than 100 meters through the entire flight.

The major difference in the fourth test was that the rhombic antenna was used. This flight terminated near Williamsburg, Mich. From this we can conclude that nearly two-thirds of the flight took place over the lake. The rhombic antenna was much bulkier than the Yagi, and matching it to the altimeter indoors was more complicated. Immediately after the launch, we could see by monitoring the RF pulses that it was not matched properly. As a result, the RF pulse was too wide and the sweeping rate too fast. Both have the overall effect of reducing the altimeter sensitivity. Despite this, and probably due to the higher gain of the antenna, meaningful height measurements were received from the thirteenth minute until the 65th minute of the flight, when the altitude was 16.8 km. In Fig. 20 we give the readings received from this flight during a period of about one minute near the 200 mbar level. Contrary to the previous flight, the altimeter did not function at all during the descent.

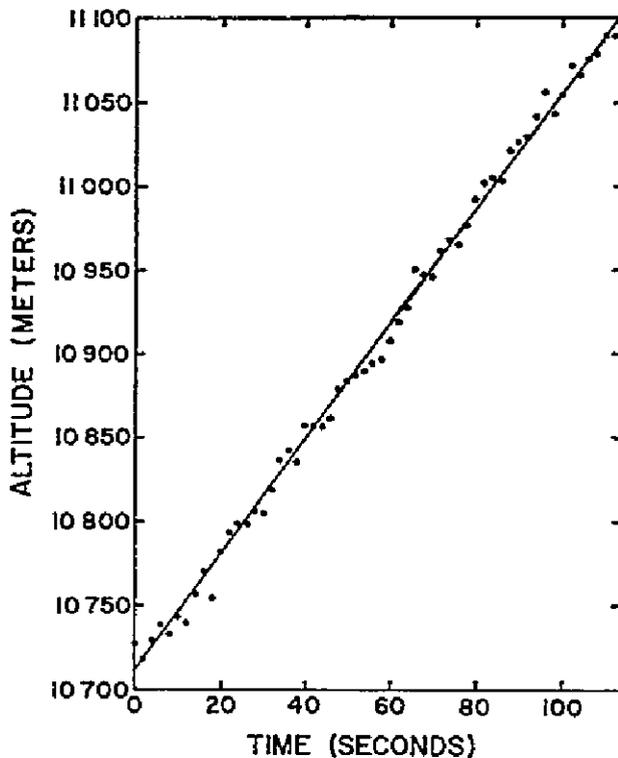


Fig. 18. Sample of the radio altimeter height readings over a two minute period during the ascent of March 14, 1969, Sturgeon Bay, Wis.

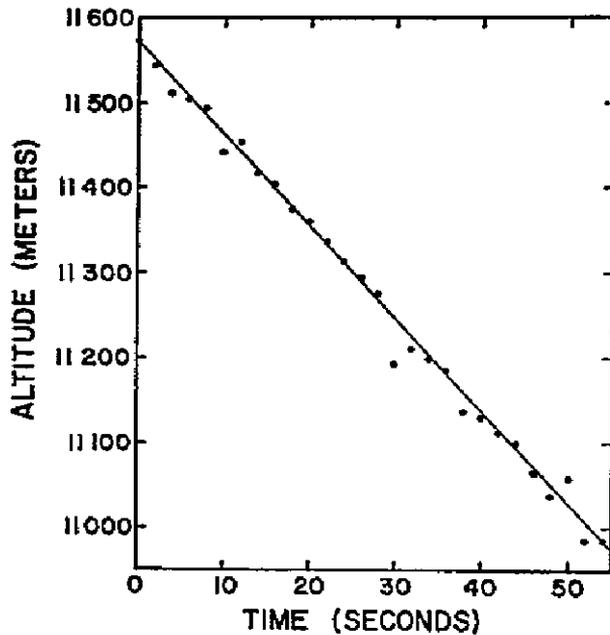


Fig. 19. Sample of the radio altimeter height readings over a one minute period during the descent of March 14, 1969, Sturgeon Bay, Wis.

A possible explanation for this could be that the rhombic antenna became tangled during the descent. Both this and the problems involved in preflight matching of the rhombic lead to the conclusion that the Yagi is more suitable. An addition of several directors would give it the necessary gain to achieve the required range of 30 km.

Summary and Conclusions

This paper has described the basic theory and design of an accurate pulse radar altimeter. This altimeter is simple and light enough to be carried aloft by a regular sounding balloon.

Altitudes up to 20 km were measured in balloon flight tests over Lake Michigan. At midrange the rms value of the random error was smaller than 0.07 percent. Height readings were available every 2 seconds during both the ascent and the descent of the balloon payload. Small improvements in the antenna would easily increase the operational range to 30 km.

When flown together with a radiosonde, this altimeter can improve the accuracy of atmospheric sounding. It can also replace either temperature or pressure measurements through the use of the hydrostatic equation. This application is most attractive for high altitude radiosondes where the pressure

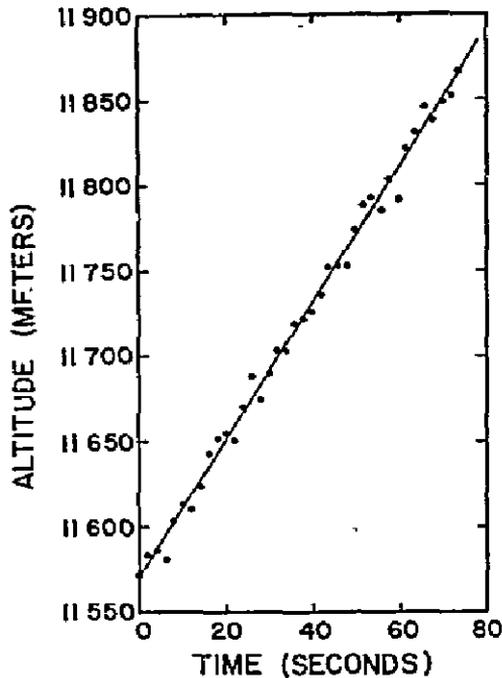


Fig. 20. Sample of the radio altimeter height readings over about a one minute period during the ascent of March 15, 1960, Sturgeon Bay, Wis.

resolution is inadequate, and for dropsondes when the thermistor time constant becomes too long compared to the rate of descent. Together with a pressure measurement, this altimeter can supply the reference pressure which is essential in the proposed system of floating superpressure balloons.

In addition to accurate height measurements, the altimeter can lock to an artificial return rather than its own returning pulse and, thus, measure distance to a transceiver. This can add wind measurement to its possible applications. All these applications should be studied and developed.

Further improvements in circuitry should also be considered. These may include: solid state RF state at 1680 MHz, combining the several ICs into one package, and elimination of all components that do not meet the design goals for air safety consideration.

Acknowledgment

I wish to extend sincere thanks to Prof. F. G. Stremmer for his advice and encouragement during the work on this research. Very special thanks are due

to Prof. V. E. Suomi, who brought me from Israel, suggested this research topic, gave me the facilities of his Space Science and Engineering Center, and was himself an endless source of ideas and inspiration. Thanks go also to C. Blair for his fine work on constructing the various models of the radio altimeter, and for his many suggestions concerning circuitry design. The help of J. Maynard in performing the flight tests was especially valuable.

References

1. G. O. Hall, "Superregenerative receivers," in Microwave Receivers, M.I.T. Rad. Lab. Ser. New York: McGraw-Hill, 1948, ch. 20.
2. V. W. Eveleigh, Adaptive Control and Optimization Techniques. New York: McGraw-Hill, 1967, ch. 9.
3. A. R. Edison, R. K. Moore, and B. D. Warner, "Radar terrain return measured at near-vertical incidence," IRE Trans. Antennas and Propagation, vol. AP-8, pp. 246-254, May 1960.
4. R. K. Moore and C. S. Williams, Jr., "Radar terrain return at near-vertical incidence," Proc. IRE, vol. 45, pp. 228-238, February 1957.
5. M. Schwartz, "Effects of signal fluctuation on signal detectability," in Modern Radar, R. S. Berkowitz, Ed. New York: Wiley, 1965, pp. 130-140.
6. A. J. Viterbi, Principles of Coherent Communication. New York: McGraw-Hill, 1966, chs. 1-4.
7. F. M. Gardner, Phaselock Techniques. New York: Wiley, 1966.
8. D. K. Barton, Radar System Analysis. Englewood Cliffs, N. J.: Prentice-Hall, 1964, pp. 25-29.
9. N. Levanon, "Balloon-borne radio altimeter," Ph.D. dissertation (Appendix B), University of Wisconsin, Madison, April 1969.

N71-11625

WOBBLE-SPIN TECHNIQUE FOR SPACECRAFT INVERSION AND
EARTH PHOTOGRAPHY

Norman H. Beachley* and John J. Uicker Jr.*

ABSTRACT

If certain prescribed spacecraft moment-of-inertia relationships are maintained, a small cylindrical flywheel can be used to tilt a spin-stabilized satellite with respect to its fixed angular momentum vector without inducing nutation and without changing the basic spin rate. A camera rigidly attached to such a tilting vehicle will have the same view as a pivoting camera in a satellite whose spin axis is fixed. This technique can be extended, if desired, to turn the satellite upside down—a very useful capability for spin-scan camera satellites using radiation-cooled instruments. The control flywheel, which can be located anywhere in the satellite, must have its axis of rotation perpendicular to the satellite spin axis. The total vehicle must have a cylindrical moment of inertia configuration, with its axis of symmetry also perpendicular to the vehicle spin axis, and parallel to the flywheel's axis of rotation.

Introduction

The present ATS satellite is being used to take high-quality pictures of the earth from a synchronous orbit. This spin-stabilized satellite rotates at 100 rpm, and photographs a single line across the earth during each revolution. After each sweep, the camera is pivoted through a small angle to allow a new line on the earth to be photographed the next time around. A total of 2400 parallel lines is used to produce a picture of the entire earth.

An alternate method to accomplish the same result is presented in this

Received August 5, 1968; revision received November 25, 1968. This work was supported in part by NASA.

*Assistant Professor of Mechanical Engineering, The University of Wisconsin, Madison.

paper.* This proposed technique is based on changing the spin axis of the satellite, thus allowing the use of a camera with fixed geometry relative to the vehicle. The maneuver is accomplished by a rotating flywheel inside the satellite, so that the momentum vector remains fixed in space. By using a system that does not induce nutation, the original spin axis will then "wobble" about the momentum vector, tracing out cones in space. The result, in terms of what the camera sees, can be shown to be identical to that obtained with a rotating camera (see Fig. 1 for a comparison of the two techniques).

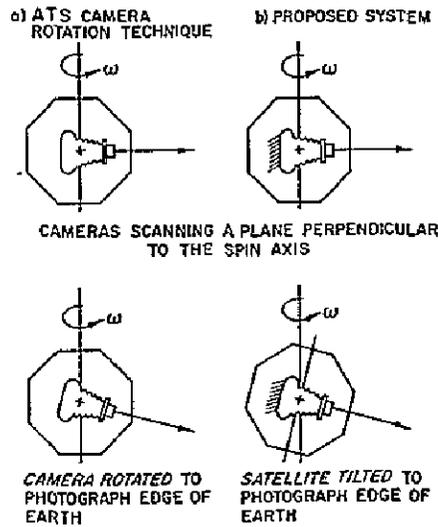


Fig. 1. Comparison of the proposed wobble-spin technique with the present ATS spin-scan system.

This method of control, allowing maneuverability without the necessity of changing the angular momentum vector, can eliminate some of the mechanical complexity of spin-scan camera satellites. This should be particularly significant in some of the systems that have been proposed for the future. A camera that can be rotated mechanically through very small increments in a space environment is a relatively complicated and expensive device and may require moving seals which introduce a reliability problem. Such cameras have been built and used with completely satisfactory results on the ATS satellites, but the problem tends to become accentuated with larger and more complex cameras. For example, an infrared camera presently under consideration would require a very large pivoting reflector mirror if used in a satellite with a fixed spin axis.

*The incentive for this work was provided by V. Suomi, University of Wisconsin, who pointed out the potential usefulness of this concept if someone could conceive a practical method of control.

Its size dictates a very rigid and probably heavy support structure to prevent damage during the launch environment. With the technique presented here, however, not only the tilting mechanism but the mirror itself could be eliminated.

Another important feature of this technique is that it will allow the spacecraft to be turned upside down with very little expenditure of energy, and without the necessity of an external torque, as would be required if the momentum vector itself were to be inverted. The only difference between rotating the satellite 180° with respect to its fixed angular momentum vector and inverting the momentum vector itself is that they result in opposite directions of rotation. This, however, should add no significant difficulty to data acquisition and utilization.

This inversion technique would be extremely valuable for any satellite whose spin axis is parallel to that of the earth, and in which there is an instrument which must be kept at a very low temperature by radiation cooling in order to work satisfactorily (such as would be the case with a spin-scan satellite with an infrared camera using a Hg-Cd-Te detector). Figure 2 shows the geometric relationship existing between the sun and such a satellite in the summer and winter, and the usefulness of satellite inversion when a radiation cooler is employed. It is apparent that the cooler should "look" along the spin axis. This gives two possible directions for it to be pointed. The direction that is correct for the spring and summer (i. e., the direction that does not allow the sun's rays to enter) is incorrect for the fall and winter, and the cooler, even if of a complex shielded design, would be less effective then. If, however, the satellite were inverted twice a year, at the vernal and autumnal equinoxes, the cooler would be able to operate at a maximum effectiveness over the entire

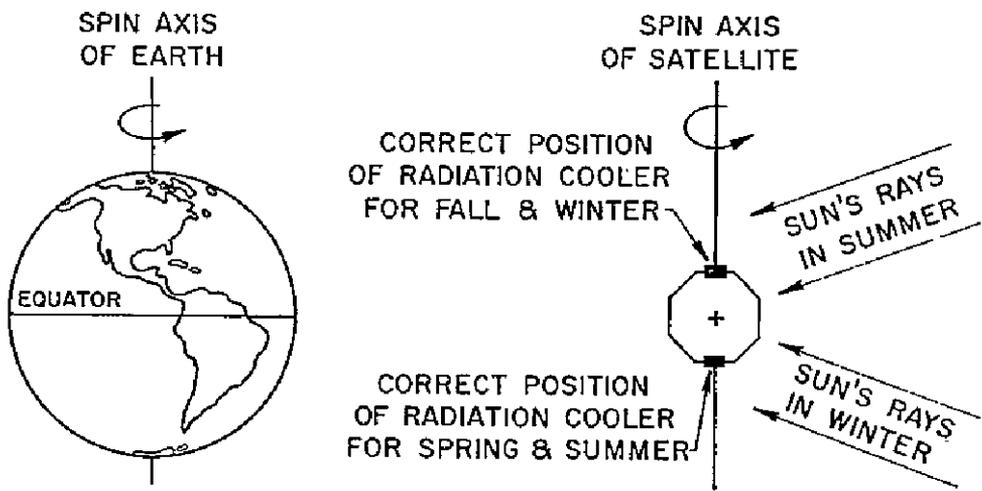


Fig. 2. Illustration of the advantage of inversion for a satellite using a radiation cooler.

Details of Proposed System

In order to accomplish this satellite tilting satisfactorily, the following requirements must be satisfied: 1) the method of control must allow the momentum vector to remain fixed in space (this is satisfied if control is accomplished by internal mass movement only); 2) the spin rate of the satellite must remain constant throughout the tilting maneuver; 3) the spin axis of the satellite must coincide with a principal axis throughout the entire maneuver, to avoid inducing unwanted nutation; 4) the control mechanism must be small, reliable, and consume little power.

A mathematical analysis of the problem (see Appendix) has led to a very simple concept that meets all of the preceding requirements. The proposed system consists of 1) the basic vehicle, which must have certain prescribed moment of inertia relationships and 2) a motor-driven cylindrical flywheel whose axis of rotation is perpendicular to the vehicle spin axis. Figure 3 is an illustration of the proposed system. e_1 , e_2 , and e_3 are principal axes of the entire satellite (including the flywheel) with the spin axis in the e_2e_3 plane. The satellite can, of course, have any shape. It is shown as a cylinder in the figure to emphasize the required moment of inertia relationships,

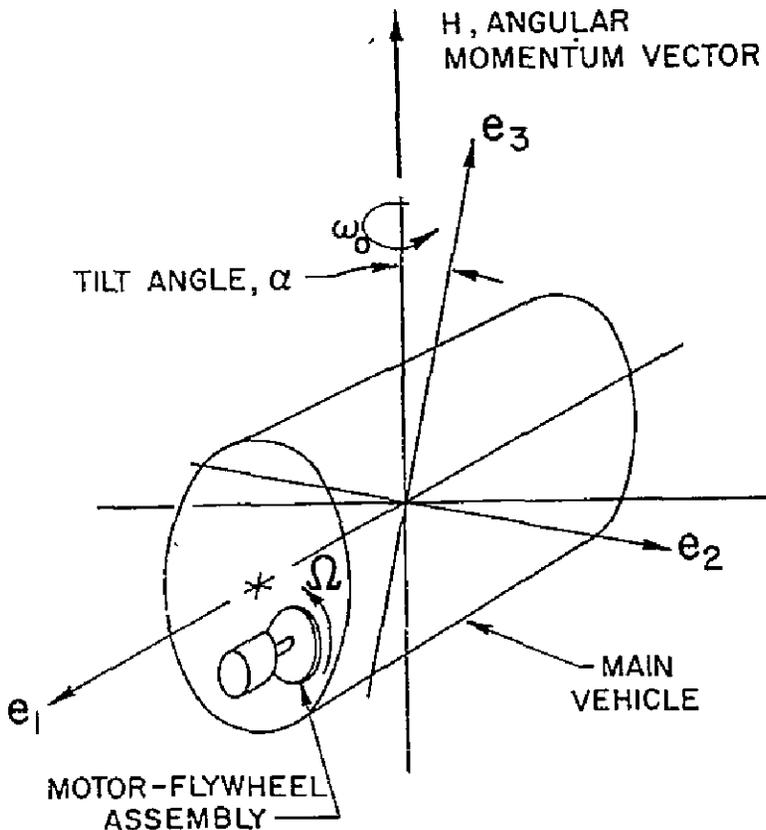


Fig. 3. Proposed wobble-spin system.

$I_{22} = I_{33}$ and $I_{22} > I_{11}$, with these values including the flywheel. With these conditions fulfilled, the flywheel may be placed anywhere in the satellite. Its axis of rotation, however, must be parallel to the vehicle e_1 axis.

To obtain spin-scan pictures with this type of system, the satellite must first be placed into synchronous orbit, and set into rotation with its spin axis parallel to that of the earth, the flywheel being at rest during these preliminary operations. The initial spin axis should be chosen so that the view of the camera lies just outside the edge of the earth. To take a picture, the flywheel is made to rotate at a constant speed, causing the basic vehicle to slowly rotate, or tilt, in the opposite direction about its e_1 axis (the basic spin action being unaffected), allowing a series of properly spaced parallel lines across the earth to be photographed. When the picture is completed, the flywheel can be rotated in the opposite direction to return the satellite to its starting position. This return time can be shortened, if desired, by running the flywheel at a higher rate of speed.

The purpose of the "cylindrical" configuration is readily demonstrated. The vehicle e_1 axis can be shown to remain perpendicular to the momentum vector throughout the tilting maneuver. With the cylindrical symmetry ($I_{22} = I_{33}$), the moments of inertia about all satellite axes passing through the center of mass and perpendicular to the e_1 axis are equal, so that any one can be considered a principal axis. Therefore, the momentum vector will always lie on a principal axis, meeting requirement 3, and the spin rate will remain constant, due to the fixed moment of inertia about the momentum vector, satisfying requirement 2.

For spin stability, it is desirable to keep the moment of inertia of the total satellite about the spin axis greater than that about the e_1 axis. This is the reason for the requirement that $I_{22} > I_{11}$.

As the flywheel is rotated, the satellite will tilt with respect to the fixed angular momentum vector, about the vehicle e_1 axis. The angle of tilt, as shown in Fig. 3, is designated as α . Analysis (see Appendix) shows that the rate of change of the tilt angle is directly proportional to the flywheel speed,

$$\dot{\alpha} = -J_0 \Omega / I_{11}$$

where J_0 is the moment of inertia of the flywheel about its axis of rotation.

It is interesting to consider the gyroscopic torques that are present in this system. With the vehicle e_1 axis perpendicular to the angular momentum vector, and the flywheel speed zero before the tilting maneuver begins, there is obviously no angular momentum about the e_1 axis. During the subsequent tilting operation, therefore, the net angular momentum about the e_1 axis must remain zero. This means that the e_1 axis angular momentum of the spinning flywheel is equal and opposite to that of the slowly rotating main vehicle. The

gyroscopic torques (equal to the product of the basic satellite spin rate and the e_1 axis angular momentum) are therefore also equal and opposite for the two bodies (flywheel and main vehicle) and occur as merely radial loads on the flywheel bearings.

As an example of the size flywheel required for this type of system, consider a 1000 lb satellite spinning at 100 rpm, with a radius of gyration of 24 in. about the e_1 axis, photographing the earth with a series of 2400 lines, the earth extending over a 15° field of view with respect to the satellite. The required tilt rate of such a satellite would be 0.625 deg/minute, or 0.00174 rpm. If this were to be caused by rotating the flywheel at, say, 1740 rpm (a very conservative figure), the required relationship between its moment of inertia and that of the main vehicle about the e_1 axis would be

$$J_0/I_{11} = -\dot{\alpha}/\Omega = 0.00174/1740.0 = 1.0 \times 10^{-6}$$

If the flywheel were made of steel 1/4 in. thick, a diameter of 2.10 in. would be required.

The motor flywheel assembly could be hermetically sealed as a unit. This should result in excellent reliability. The power requirements of a motor to drive such a flywheel should be small.

APPENDIX: Equations of Motion

Nomenclature

- a = acceleration of the origin of the e_i coordinate system
- e_i = unit vectors of a right-hand Cartesian coordinate system fixed in the main body, but with origin at the center of mass of the total satellite and directed along the principal axes of inertia of the total satellite
- H_1 = angular momentum vector of the main body
- H_2 = angular momentum vector of the flywheel
- I_{ii} = moments of inertia of the total satellite about its principal axes e_i
- $J_{ij}^{(1)}$ = moments and products of inertia of the main body about a set of axes parallel to e_i but passing through the center of mass of the main body
- $J_{ij}^{(2)}$ = moments and products of inertia of the flywheel about a set of axes instantaneously parallel to e_i but passing through its center of mass

- J_0 = moment of inertia of the flywheel about its axis of rotation
 m_1 = mass of the main body
 m_2 = mass of the flywheel
 s = first moment of mass of the satellite about the origin of the e_1 coordinate system
 α = angle between the angular momentum vector of the satellite and the e_3 axis
 ρ_1 = position vector of the center of mass of the main body with respect to the origin of the e_1 coordinate system
 ρ_2 = position vector of the center of mass of the flywheel with respect to the origin of the e_1 coordinate system
 Ω, Ω_i, Ω = angular velocity vector of the flywheel relative to the main body (a known function of time), components of Ω along the e_i axes, magnitude of the vector Ω
 ω, ω_i = angular velocity of the main body, components of ω along the e_i axes
 ω_0 = initial spin rate of the satellite

Assumptions

1) The satellite under study consists of two rigid masses, a main body and a flywheel, which are pivotally connected together and which can be rotated relative to one another by a controlled source of power.

2) The flywheel, which may be located anywhere in the main body, is a cylinder which rotates relative to the main body about its axis of symmetry. The moments of inertia of the total satellite are therefore not affected by the flywheel's rotation.

3) The net external force and torque acting on the satellite are both negligible.

The generalized expression for the conservation of angular momentum for the type of system being considered is [1]

$$\dot{H}_1 + \dot{H}_2 + s \times a = 0 \quad (A1)$$

where H_1 denotes the angular momentum of the main body and H_2 that of the flywheel. Both terms are calculated with respect to an arbitrary reference point whose linear acceleration is a , and s represents the first moment of mass of the system about that point.

By choosing the reference point at the center of mass of the total satellite,

$$s \times a = 0 \quad (A2)$$

If this point is also chosen as the origin of a coordinate system fixed in the main body, then the angular momentum of the main body is

$$H_1 = e_i J_{ij}^{(1)} \omega_j + m_1 \rho_1 \times \dot{\rho}_1 \quad (i, j = 1, 2, 3) \quad (A3)$$

With the vector Ω defined as the angular velocity of the flywheel relative to the e_i coordinate system, the angular momentum of the flywheel is

$$H_2 = e_i J_{ij}^{(2)} (\omega_j + \Omega_j) + m_2 \rho_2 \times \dot{\rho}_2 \quad (i, j = 1, 2, 3) \quad (A4)$$

The total angular momentum of the satellite may now be written

$$H_1 + H_2 = [e_i (J_{ij}^{(1)} + J_{ij}^{(2)}) \omega_j + m_1 \rho_1 \times \dot{\rho}_1 + m_2 \rho_2 \times \dot{\rho}_2] + e_i J_{ij}^{(2)} \Omega_j \quad (i, j = 1, 2, 3) \quad (A5)$$

If there were no relative motion between the flywheel and main body and if the angular velocity vector of the resultant rigid body were ω , then its angular momentum would be given by the bracketed terms of Eq. (A5). With the coordinate axes e_i chosen to be directed along the principal axes of inertia of the total satellite, however, an alternate expression for the angular momentum of such a rigid body is $e_i I_{ii} \omega_i$. Therefore, Eq. (A5) can be written in the simpler form

$$H_1 + H_2 = e_i I_{ii} \omega_i + e_i J_{ij}^{(2)} \Omega_j \quad (i, j = 1, 2, 3) \quad (A6)$$

If we now add the additional restriction that the flywheel's axis of rotation relative to the main body is parallel to the e_1 axis, the angular momentum of the satellite reduces to

$$H_1 + H_2 = (I_{11} \omega_1 + J_0 \Omega) e_1 + I_{22} \omega_2 e_2 + I_{33} \omega_3 e_3 \quad (A7)$$

Substituting this expression into Eq. (A1) and using Eq. (A2), we see that

$$(I_{11} \dot{\omega}_1 + J_0 \dot{\Omega}) e_1 + I_{22} \dot{\omega}_2 e_2 + I_{33} \dot{\omega}_3 e_3 + (I_{11} \omega_1 + J_0 \Omega) \dot{e}_1 + I_{22} \omega_2 \dot{e}_2 + I_{33} \omega_3 \dot{e}_3 = 0 \quad (A8)$$

Since the e_i coordinate axes are fixed in the main body, the time rates of change of the unit vectors are given by

$$\dot{e}_1 = \omega \times e_1 = \omega_3 e_2 - \omega_2 e_3 \quad (A9a)$$

$$\dot{e}_2 = \omega \times e_2 = \omega_1 e_3 - \omega_3 e_1 \quad (A9b)$$

$$\dot{e}_3 = \omega \times e_3 = \omega_2 e_1 - \omega_1 e_2 \quad (A9c)$$

Substituting these expressions into Eq. (A8), we obtain

$$\begin{aligned} & [I_{11}\dot{\omega}_1 + J_0\dot{\Omega} + (I_{33} - I_{22})\omega_2\omega_3]e_1 + \\ & [I_{22}\dot{\omega}_2 + J_0\Omega\omega_3 + (I_{11} - I_{33})\omega_1\omega_3]e_2 + \\ & [I_{33}\dot{\omega}_3 - J_0\Omega\omega_2 - (I_{11} - I_{22})\omega_1\omega_2]e_3 = 0 \end{aligned} \quad (A10)$$

The governing differential equations of motion for the satellite are now found by equating each component of Eq. (A10) to zero:

$$I_{11}\dot{\omega}_1 = -J_0\dot{\Omega} + (I_{22} - I_{33})\omega_2\omega_3 \quad (A11)$$

$$I_{22}\dot{\omega}_2 = -J_0\Omega\omega_3 - (I_{11} - I_{33})\omega_1\omega_3 \quad (A12)$$

$$I_{33}\dot{\omega}_3 = J_0\Omega\omega_2 + (I_{11} - I_{22})\omega_1\omega_2 \quad (A13)$$

Since the differential equations are nonlinear and cross coupled, their solution in closed form is not apparent. Equation (A11) can be uncoupled, however, by placing an additional restriction on the configuration of the total satellite, i. e., by requiring that its mass moments of inertia about the e_2 and e_3 principal axes be equal,

$$I_{22} = I_{33} \quad (A14)$$

With this restriction, the equations of motion become

$$I_{11}\dot{\omega}_1 = -J_0\dot{\Omega} \quad (A15)$$

$$I_{22}\dot{\omega}_2 = -J_0\Omega\omega_3 - (I_{11} - I_{22})\omega_1\omega_3 \quad (A16)$$

$$I_{22}\dot{\omega}_3 = J_0\Omega\omega_2 + (I_{11} - I_{22})\omega_1\omega_2 \quad (A17)$$

Equation (A15) can now be readily integrated to give

$$\omega_1 = -J_0\Omega/I_{11} \quad (A18)$$

with boundary conditions assumed such that $\omega_1 = 0$ when $\Omega = 0$. Using this result, Eqs. (A16) and (A17) reduce to

$$\dot{\omega}_2 = -J_0 \Omega \omega_3 / I_{11} \quad (\text{A19})$$

$$\dot{\omega}_3 = J_0 \Omega \omega_2 / I_{11} \quad (\text{A20})$$

Equation (A18) shows that ω_1 is a function of Ω only and, since it was assumed that $\omega_1 = 0$ when $\Omega = 0$, the angular momentum vector of the system when $\Omega = 0$ must be in the $e_2 e_3$ plane. Since ω_1 is not a function of ω_2 or ω_3 , it will never include any component of the vehicle's original spin rate, and the angle of tilt of the main body from its original position must take place in the $e_2 e_3$ plane. Therefore, regardless of the angle α between the angular momentum vector of the total satellite and the e_3 axis, the following relationships must hold:

$$\omega_2 = \omega_0 \sin \alpha \quad (\text{A21})$$

$$\omega_3 = \omega_0 \cos \alpha \quad (\text{A22})$$

where ω_0 is the satellite's initial (and constant, since $I_{33} = I_{22}$) spin rate about the angular momentum vector.

Substituting these expressions into Eq. (A19) or (A20) gives the equation relating the rate of tilt of the vehicle to the speed of the flywheel,

$$\dot{\alpha} = -J_0 \Omega / I_{11} \quad (\text{A23})$$

Reference

1. Grubin, C., "On Generalization of the Angular Momentum Equation," Journal of Engineering Education, Vol. 51, 1960, p. 237.

